ENVS C100: ENVIRONMENTAL SCIENCE



Tanya Hoerer
Coastline College

LibreTexts™

ENVS C100: Environmental Science (Hoerer)

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

This text was compiled on 05/03/2026

TABLE OF CONTENTS

Licensing

1: Introduction

- 1.1: Environmental Science Preface
- 1.2: What Is Science, and How Does It Work?
  - 1.2.1: What is Science?
  - 1.2.2: The Scientific Method
  - 1.2.3: Scientific Papers
  - 1.2.4: Basic and Applied Science
  - 1.2.5: Data Dive- Tonsillectomy Trends
  - 1.2.6: Review
- 1.3: Math Blast- An Overview of Essential Mathematics Used in Science
  - 1.3.1: Getting down the basics- What are the essential mathematical and statistical terms and concepts?
  - 1.3.2: Summarizing the data- Descriptive statistics
  - 1.3.3: Putting it all together- Inferential statistics and hypothesis testing
  - 1.3.4: Data Dive- California Wildfires
  - 1.3.5: Review

2: Environmental Chemistry

- 00: Front Matter
  - TitlePage
  - InfoPage
  - Table of Contents
  - Licensing
- 2.1: Geochemistry (Lower)
  - Licensing
  - 2.1.1: The Earth and its Lithosphere
    - 2.1.1.1: Composition and Structure of the Earth
    - 2.1.1.2: Origin of the Elements
    - 2.1.1.3: Formation and evolution of the Earth
    - 2.1.1.4: The Earth's crust
  - 2.1.2: The Hydrosphere
    - 2.1.2.1: Water, Water Everywhere...
    - 2.1.2.2: The hydrosphere and the oceans
    - 2.1.2.3: Chemistry and geochemistry of the oceans
    - 2.1.2.4: Chemical budgets of oceanic elements
  - 2.1.3: The Atmosphere
    - 2.1.3.1: Structure and Composition of the Atmosphere
    - 2.1.3.2: Origin and Evolution of the Atmosphere
  - 2.1.4: The Biosphere
    - 2.1.4.1: Chemistry and Energetics of the Life Process
    - 2.1.4.2: Biogeochemical Evolution

- 2.1.4.3: Gaia - Bioregulation of the Environment
- 2.2: Supplemental Modules (Environmental Chemistry)
  - 2.2.1: Acid Rain
    - 2.2.1.1: Acid Rain
    - 2.2.1.2: Acid Rain Transport
    - 2.2.1.3: Acid Snow
    - 2.2.1.4: Electricity Generation
    - 2.2.1.5: Playground 1
    - 2.2.1.6: Sources of Nitrogen Oxides
    - 2.2.1.7: Sources of Sulfur Oxides
  - 2.2.2: Aquatic Chemistry
    - 2.2.2.1: 1. Acid-Base Chemistry of Natural Aquatic Systems
    - 2.2.2.2: 2. Carbonate Equilibria in Natural Waters
    - 2.2.2.3: 3. Redox Equilibria in Natural Waters
    - 2.2.2.4: 4. Solids in Contact With Natural Waters
    - 2.2.2.5: Fundamental Characteristics of Water
    - 2.2.2.6: Natural Water
    - 2.2.2.7: Water Biology
    - 2.2.2.8: Water Chemistry
    - 2.2.2.9: Water Physics
    - 2.2.2.10: Water Treatment
  - 2.2.3: Atmospheric Chemistry
    - 2.2.3.1: Atmosphere
    - 2.2.3.2: Carbon Cycle
    - 2.2.3.3: Carbon Oxides
    - 2.2.3.4: Ozone
    - 2.2.3.5: Photosynthesis
  - 2.2.4: Biochemical Cycles
    - 2.2.4.1: Biochemical Cycles
    - 2.2.4.2: Carbon Cycle
    - 2.2.4.3: Nitrogen Cycle
- 2.3: Environmental Toxicology (van Gestel et al.)
  - 00: Front Matter
    - Table of Contents
  - 2.3.1: Environmental Toxicology
    - 2.3.1.1: Environmental toxicology
    - 2.3.1.2: DPSIR
    - 2.3.1.3: Short history
  - 2.3.2: Environmental Chemistry, Chemicals
    - 2.3.2.1: Introduction
    - 2.3.2.2: Pollutants with specific properties
    - 2.3.2.3: Pollutants with specific use
  - 2.3.3: Environmental Chemistry - From Fate to Exposure
    - 2.3.3.1: Environmental compartments
    - 2.3.3.2: Sources of chemicals
    - 2.3.3.3: Pathways and processes determining chemical fate
    - 2.3.3.4: Partitioning and Partitioning Constants

- 2.3.3.5: Metal Speciation
- 2.3.4: Toxicology
  - 2.3.4.1: Toxicokinetics
  - 2.3.4.2: Toxicodynamics and Molecular Interactions
  - 2.3.4.3: Toxicity Testing
  - 2.3.4.4: Increasing ecological realism in toxicity testing
  - 2.3.4.5: Availability and bioavailability
  - 2.3.4.6: Degradation
  - 2.3.4.7: Modelling exposure
  - 2.3.4.8: Section 4.1. Toxicokinetics
  - 2.3.4.9: Section 4.4. Increasing ecological realism in toxicity testing
  - 2.3.4.10: Section 4.2. Toxicodynamics and Molecular Interactions
  - 2.3.4.11: Section 4.3. Toxicity testing
- 2.3.5: Population, Community and Ecosystem Ecotoxicology
  - 2.3.5.1: Introduction- Linking population, community and ecosystem responses
  - 2.3.5.2: Population ecotoxicology in laboratory settings
  - 2.3.5.3: Wildlife population ecotoxicology
  - 2.3.5.4: Trait-based approaches
  - 2.3.5.5: Population models
  - 2.3.5.6: Metapopulations
  - 2.3.5.7: Community ecotoxicology
  - 2.3.5.8: Structure versus function incl. ecosystem services
  - 2.3.5.9: Landscape ecotoxicology
- 2.3.6: Risk Assessment and Regulation
  - 2.3.6.1: Introduction - The Essence of Risk Assessment
  - 2.3.6.2: Ecosystem services and protection goals
  - 2.3.6.3: Predictive risk assessment approaches and tools
  - 2.3.6.4: Diagnostic risk assessment approaches and tools
  - 2.3.6.5: Regulatory Frameworks
  - 2.3.6.6: Risk management and risk communication
  - 2.3.6.7: Risk perception
- 2.4: Key Elements of Green Chemistry (Lucia)
  - Licensing
  - 2.4.1: Key Elements of Green Chemistry
    - 2.4.1.1: Introduction to Green Chemistry
    - 2.4.1.2: Green Chemistry Concepts
    - 2.4.1.3: Case Study
    - 2.4.1.4: Conclusions and Review
  - 2.4.2: Life-Cycle Analysis
    - 2.4.2.1: Introduction to Life Cycles
    - 2.4.2.2: LCA/LCIA Concepts
    - 2.4.2.3: Conclusions and Review
  - 2.4.3: Hazards
    - 2.4.3.1: Introduction to Hazards of Chemistry
    - 2.4.3.2: Hazard Concepts
    - 2.4.3.3: Case Study - Badger Army Ammunition Plant
    - 2.4.3.4: Green Technologies for Safer Chemical Production

- 2.4.3.5: Conclusions and Review Questions
- 2.4.4: Alternative Solvents
  - 2.4.4.1: Introduction to Solvents
  - 2.4.4.2: Solvent Concepts
  - 2.4.4.3: Solvent Alternatives
  - 2.4.4.4: Conclusions and Review
- 2.4.5: Alternative Reagants
  - 2.4.5.1: Introduction to Reagants
  - 2.4.5.2: Reagant Concepts
  - 2.4.5.3: Specific Replacements
  - 2.4.5.4: Conclusions and Review
- 2.4.6: Reaction Types, Design, and Efficiency
  - 2.4.6.1: Introduction to Reactions
  - 2.4.6.2: Reaction Concepts
  - 2.4.6.3: Reaction Design Concepts
  - 2.4.6.4: Conclusions and Review
- Detailed Licensing

- 2.5: Green Chemistry and the Ten Commandments of Sustainability (Manahan)
  - Licensing
  - 2.5.1: Sustainability and the Environment
    - 2.5.1.1: Sustainability
    - 2.5.1.2: The Environment and the Five Environmental Sphere
    - 2.5.1.3: Seeing Green
    - 2.5.1.4: Natural Capital of the Earth
    - 2.5.1.5: Sustainability as a Group Effort- It Takes a (Very Big) Village
    - 2.5.1.6: Sustainable Energy- Away from the Sun and Back Again
    - 2.5.1.7: Green Science
    - 2.5.1.8: Green Technology
    - 2.5.1.9: Sustainability and the Eco-Economy
    - 2.5.1.10: Green Products and Services- Design for Sustainability
    - 2.5.1.11: Questions and Problems
  - 2.5.2: The Key Role of Chemistry and Making Chemistry Green
    - 2.5.2.1: Chemistry is Good (and Unavoidable)
    - 2.5.2.2: The Environment and its Chemistry
    - 2.5.2.3: What is Environmental Chemistry?
    - 2.5.2.4: Environmental Pollution
    - 2.5.2.5: Practice of Green Chemistry
    - 2.5.2.6: Green Chemistry and Synthetic Chemistry
    - 2.5.2.7: Reduction of Risk- Hazard and Exposure
    - 2.5.2.8: The Risks of No Risks
    - 2.5.2.9: Waste Prevention
    - 2.5.2.10: Basic Principles of Green Chemistry
    - 2.5.2.11: Some Things to Know About Chemistry before You Even Start
    - 2.5.2.12: Combining Atoms to Make Molecules and Compounds
    - 2.5.2.13: The process of making and breaking chemical bonds- chemical reactions
    - 2.5.2.14: The Nature of Matter and States of Matter
    - 2.5.2.15: References and Questions

- 2.5.3: The Elements - Basic Building Blocks of Green Chemicals
  - 2.5.3.1: Elements, Atoms, and Atomic Theory
  - 2.5.3.2: Hydrogen - The Simplest Atom
  - 2.5.3.3: Helium - The First Noble Gas
  - 2.5.3.4: Lithium, The First Metal
  - 2.5.3.5: The Second Period of the Periodic Table
  - 2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons
  - 2.5.3.7: Completing the 20-Element Periodic Table
  - 2.5.3.8: The Brief Periodic Table is Complete
  - 2.5.3.E: The Elements - Basic Building Blocks of Green Chemicals (Exercises)
  - 2.5.3.R: References
- 2.5.4: Compounds- Safer Materials for a Safer World
  - 2.5.4.1: Chemical Bonds and Compound Formation
  - 2.5.4.2: Electrons Involved in Chemical Bonds and Octets of Electrons
  - 2.5.4.3: Sodium Chloride and Ionic Bonds
  - 2.5.4.4: Covalent Bonds in H2 and Other Molecules
  - 2.5.4.5: Covalent Bonds in Compounds
  - 2.5.4.6: Covalent Bonds and Green Chemistry
  - 2.5.4.7: Predicting Formulas of Covalently Bound Compounds
  - 2.5.4.8: Chemical Formulas, the Mole, and Percentage Composition
  - 2.5.4.9: What Are Chemical Compounds Called?
  - 2.5.4.10: Acids, Bases and Salts
  - 2.5.4.11: Questions and Problems
- 2.5.5: Chemical Reactions- Making Materials Safely and Sustainable
  - 2.5.5.1: Describing What Happens with Chemical Equations
  - 2.5.5.2: Balancing Chemical Equations
  - 2.5.5.3: Just Because You Can Write It Doesn't Mean That It Will Happen
  - 2.5.5.4: Yield And Atom Economy in Chemical Reactions
  - 2.5.5.5: Catalysts That Make Reactions Go
  - 2.5.5.6: Kinds of Chemical Reactions
  - 2.5.5.7: Oxidation-Reduction Reactions and Green Chemistry
  - 2.5.5.8: Quantitative Information from Chemical Reactions
  - 2.5.5.9: Energy in Chemical Reactions
  - 2.5.5.10: Stoichiometry by the Mole Ratio Method
  - 2.5.5.11: Limiting Reactant and Percent Yield
  - 2.5.5.12: Titrations - Measuring Moles by Volume of Solution
  - 2.5.5.13: Industrial Chemical Reactions - The Solvay Process
  - 2.5.5.14: Questions and Problems
- 2.5.6: The Wonderful World of Carbon - Organic Chemistry and Biochemicals
  - 2.5.6.1: Rings and Chains of Carbon Atoms
  - 2.5.6.2: Compounds of Carbon and Hydrogen - Hydrocarbons
  - 2.5.6.3: Using Lines To Show Organic Structural Formulas
  - 2.5.6.4: Functional Groups
  - 2.5.6.5: Giant Molecules from Small Organic Molecules
  - 2.5.6.6: Questions and Problems
  - 2.5.6.7: Supplementary References
- 2.5.7: Chemistry of Life and Green Chemistry
  - 2.5.7.1: Green Biochemistry

- 2.5.7.2: Biochemistry and the Cells
- 2.5.7.3: Carbohydrates
- 2.5.7.4: Proteins
- 2.5.7.5: Lipids - Fats, Oils, and Hormones
- 2.5.7.6: Nucleic Acids
- 2.5.7.7: Enzymes
- 2.5.7.8: Biochemical Processes in Metabolism
- 2.5.7.9: Biochemistry of Toxic Substances and Toxicological Chemistry
- 2.5.7.10: Literature Cited
- 2.5.7.11: Questions and Problems
- 2.5.7.12: Supplementary References

- 2.5.8: The Five Environmental Spheres and Biogeochemical Cycle
  - 2.5.8.1: Five Environmental Spheres
  - 2.5.8.2: The Hydrosphere
  - 2.5.8.3: The Atmosphere
  - 2.5.8.4: The Geosphere
  - 2.5.8.5: The Biosphere
  - 2.5.8.6: The Anthrosphere
  - 2.5.8.7: Cycles of Matter
  - 2.5.8.8: Questions and Problems
  - 2.5.8.9: Supplementary References
- 2.5.9: Water - the Ultimate Green Substance
  - 2.5.9.1: H_2O - Simple Formula, Remarkable Molecule
  - 2.5.9.2: Occurrence, Availability , and Utilization of Water on Earth
  - 2.5.9.3: Water Chemistry
  - 2.5.9.4: Water Pollution
  - 2.5.9.5: Greening of Water - Purification Before Use
  - 2.5.9.6: Wastewater Treatment
  - 2.5.9.7: Advanced Wastewater Treatment and Recycling
  - 2.5.9.8: The Many Uses of Water
  - 2.5.9.9: Hot Water - Pressurized Subcritical Water
  - 2.5.9.10: Supercritical Water
  - 2.5.9.11: Literature Cited
  - 2.5.9.12: Questions and Problems
  - 2.5.9.13: Supplementary References
- 2.5.10: Blue Skies for a Green Environment
  - 2.5.10.1: More Than Just Air to Breathe
  - 2.5.10.2: The Gas Laws
  - 2.5.10.3: The Protective Atmosphere
  - 2.5.10.4: Atmospheric Chemistry and Photochemical Reactions
  - 2.5.10.5: The Enormous Importance of Climate
  - 2.5.10.6: Atmospheric Particle Pollutants
  - 2.5.10.7: Pollutant Gaseous Oxides
  - 2.5.10.8: Acid Rain
  - 2.5.10.9: Miscellaneous Gases in the Atmosphere
  - 2.5.10.10: Photochemical Smog
  - 2.5.10.11: Natural Capital of the Atmosphere
  - 2.5.10.12: Energy and Mass Transfer in the Atmosphere
  - 2.5.10.13: Questions and Problems

- 2.5.10.14: Supplementary References
- 2.5.11: The Geosphere and a Green Earth
  - 2.5.11.1: Probing Deep into the Geosphere - A Well Too Far?
  - 2.5.11.2: The Nature of the Geosphere
  - 2.5.11.3: The Geosphere as a Source of Natural Capital
  - 2.5.11.4: Environmental Hazards of the Geosphere
  - 2.5.11.5: Water on and in the Geosphere
  - 2.5.11.6: Anthrospheric Influences on the Geosphere
  - 2.5.11.7: The Geosphere as a Waste Repository
  - 2.5.11.8: Have You Thanked a Clod Today?
  - 2.5.11.9: Plant Nutrients and Fertilizers in Soil
  - 2.5.11.10: Soil and Plants Related To Wastes and Pollutants
  - 2.5.11.11: Soil Loss - Desertification and Deforestation
  - 2.5.11.12: Literature Cited and Supplementary References
  - 2.5.11.13: Questions and Problems
- 2.5.12: The Biosphere and the Role of Green Chemistry in Feeding a Hungry World
  - 2.5.12.1: Pigweed's Revenge
  - 2.5.12.2: The Biosphere
  - 2.5.12.3: Cells - Basic Units of Life
  - 2.5.12.4: Metabolism and Control in Organisms
  - 2.5.12.5: Reproduction and Inherited Traits
  - 2.5.12.6: Stability and Equilibrium of the Biosphere
  - 2.5.12.7: DNA and the Human Genome
  - 2.5.12.8: Genetic Engineering
  - 2.5.12.9: Biological Interaction with Environmental Chemicals
  - 2.5.12.10: Biodegradation
  - 2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture
  - 2.5.12.12: Agricultural Applications of Genetically Modified Organisms
  - 2.5.12.13: The Anthrosphere in Support of the Biosphere
  - 2.5.12.14: Livestock and their Wastes
  - 2.5.12.15: Questions and Problems
  - 2.5.12.16: Supplementary References
- 2.5.13: The Anthrosphere, Industrial Ecology, and Green Chemistry
  - 2.5.13.1: Industrial Ecology and Industrial Ecosystems
  - 2.5.13.2: Metabolic Processes in Industrial Ecosystems
  - 2.5.13.3: Life Cycles in Industrial Ecosystems
  - 2.5.13.4: Kinds of Products
  - 2.5.13.5: Attributes Required by an Industrial Ecosystem
  - 2.5.13.6: The Kalundborg Industrial Ecosystem
  - 2.5.13.7: Environmental Impacts of Industrial Ecosystems
  - 2.5.13.8: Green Chemistry and Industrial Ecology
  - 2.5.13.9: Predicting and Reducing Hazards
  - 2.5.13.10: The E-Factor in Green Chemistry
  - 2.5.13.11: Catalysts and Catalysis
  - 2.5.13.12: Biocatalysis with Enzymes
  - 2.5.13.13: Energizing Chemical Reactions and Process Intensification
  - 2.5.13.14: Solvents and Alternate Reaction Media
  - 2.5.13.15: Feedstocks and Reagents
  - 2.5.13.16: Literature Cited and Supplementary References

- 2.5.13.17: Questions and Problems
- 2.5.14: Feeding the Anthrosphere- Utilizing Renewable and Biological Materials
  - 2.5.14.1: Feeding the Anthrosphere
  - 2.5.14.2: Utilization of Feedstocks
  - 2.5.14.3: Biological Feedstocks
  - 2.5.14.4: Biological Sources of Chemicals
  - 2.5.14.5: Biorefineries and Biomass Utilization
  - 2.5.14.6: Monosaccharide Feedstocks - Glucose, Fructose, and Xylose
  - 2.5.14.7: Cellulose
  - 2.5.14.8: Feedstocks from Cellulose Wastes
  - 2.5.14.9: Lignin
  - 2.5.14.10: Direct Biosynthesis of Polymers
  - 2.5.14.11: Bioconversion Processes for Synthetic Chemicals
  - 2.5.14.12: Bamboo - Ancient Material for the Future
  - 2.5.14.13: Literature Cited
  - 2.5.14.14: Questions and Problems
- 2.5.15: Sustainable Energy- The Essential Basis of Green Systems
  - 2.5.15.1: Sustainability and Energy
  - 2.5.15.2: What is Energy?
  - 2.5.15.3: Radiant Energy from the Sun
  - 2.5.15.4: Sources of Energy Used in the Anthrosphere
  - 2.5.15.5: Conversions Between Forms of Energy
  - 2.5.15.6: Green Technology for Energy Conversion
  - 2.5.15.7: Energy Conservation
  - 2.5.15.8: Depletable Fossil Fuels
  - 2.5.15.9: Carbon Sequestration for Fossil Fuel Utilization
  - 2.5.15.10: Nuclear Energy
  - 2.5.15.11: Renewable Energy Sources - Solar Energy
  - 2.5.15.12: Energy from Wind and Water
  - 2.5.15.13: Biomass Energy
  - 2.5.15.14: Geothermal Energy
  - 2.5.15.15: Hydrogen for Energy Storage and Utilization
  - 2.5.15.16: Literature Cited and Supplementary References
  - 2.5.15.17: Questions and Problems
- 2.5.16: Terrorism, Toxicity, and Vulnerability- Green Chemistry and Technology in Defense of Human Welfare
  - 2.5.16.1: Vulnerability to Terrorist Attack
  - 2.5.16.2: Protecting the Anthrosphere
  - 2.5.16.3: Substances That Explode, Burn, or React Violently
  - 2.5.16.4: Toxic Substances and Toxicology
  - 2.5.16.5: Toxic Chemical Attack
  - 2.5.16.6: Protecting Water, Food, and Air
  - 2.5.16.7: Detecting Hazards
  - 2.5.16.8: Green Chemistry to Combat Terrorism
  - 2.5.16.9: Green Chemistry for Sustainable Prosperity and a Safer World
  - 2.5.16.10: Green Chemistry to Combat Terrorism
  - 2.5.16.11: Green Chemistry for Sustainable Prosperity and and a Safer World
  - 2.5.16.12: Literature Cited
  - 2.5.16.13: Questions and Problems
- 2.5.17: The Ten Commandments of Sustainability and Sensible Measure

- 2.5.17.1: We Cannot Go on Like This
- 2.5.17.2: The First Commandment
- 2.5.17.3: The Second Commandment
- 2.5.17.4: The Third Commandment
- 2.5.17.5: The Fourth Commandment
- 2.5.17.6: The Fifth Commandment
- 2.5.17.7: The Sixth Commandment
- 2.5.17.8: The Seventh Commandment
- 2.5.17.9: The Ninth Commandment
- 2.5.17.10: The Tenth Commandment
- 2.5.17.11: Some Sensible Measures for Sustainability
- 2.5.17.12: The Eighth Commandment
- 2.5.17.13: Literature Cited
- 2.5.17.14: Questions and Problems
- 2.5.17.15: Supplementary References
- Detailed Licensing

- 2.6: Toxicology MSDT
  - Licensing
  - Preface
  - 2.6.1: Pathophysiology
    - 2.6.1.1: What is Pathophysiology?
    - 2.6.1.2: Outcomes of Targeted and Non-Targeted Toxicity
    - 2.6.1.3: Cellular Response to Toxicant-Induced Injury
    - 2.6.1.4: Repair and Adaptation
    - 2.6.1.5: Patterns of Toxic Injury
    - 2.6.1.6: Final Evaluation
  - 2.6.2: Biochemistry and Molecular Genetics
    - 2.6.2.1: Introduction to Biomolecules and Cell Components
    - 2.6.2.2: Cell Structure and Subcellular Compartments
    - 2.6.2.3: DNA and RNA Metabolism
    - 2.6.2.4: Epigenetic Mechanisms
    - 2.6.2.5: Section 2 Final Evaluation
  - 2.6.3: Principles of Genetic Toxicology
    - 2.6.3.1: Introduction to Genetic-toxicology Assay
    - 2.6.3.2: Different Genetic Damages or Mutations
    - 2.6.3.3: Different Genetic-Toxicology Assays
    - 2.6.3.4: Different Cytotoxicity Assays
    - 2.6.3.5: Epigenetics Assay
    - 2.6.3.6: Section 3 Final Evaluation
  - 2.6.4: Applied Systems Toxicology
    - 2.6.4.1: Systems Toxicology
    - 2.6.4.2: Dose Level and Applied Toxicology
    - 2.6.4.3: Tools and Technologies in Systems Toxicology
    - 2.6.4.4: Other Approaches for Predictive Toxicity Modeling
    - 2.6.4.5: Technologies Used In Systems Biology/Toxicology
    - 2.6.4.6: Takeaways Summary
    - 2.6.4.7: Section 4 Final Evaluation

- 2.6.5: Regulatory Toxicology
  - 2.6.5.1: Introduction to Regulatory Toxicology
  - 2.6.5.2: Global Regulatory Toxicology
  - 2.6.5.3: Topic 3- Regional Regulatory Toxicology
  - 2.6.5.4: National Regulatory Toxicology
  - 2.6.5.5: State Regulatory Toxicology
  - 2.6.5.6: Non-Governmental Regulatory Toxicology
  - 2.6.5.7: Section 5 Final Evaluation
- 2.6.6: Principles of Toxicology
  - 2.6.6.1: Introduction to Toxicology
    - 2.6.6.1.1: What is toxicology?
    - 2.6.6.1.2: Basic Terminology
  - 2.6.6.2: Dose and Dose Response
    - 2.6.6.2.1: Dose and It's Impact on Toxicity
    - 2.6.6.2.2: The Dose Response Relationship
    - 2.6.6.2.3: Dose Estimates of Toxic Effects
    - 2.6.6.2.4: Determining the Safety of a Drug
    - 2.6.6.2.5: NOAEL and LOAEL
  - 2.6.6.3: Toxic Effects
    - 2.6.6.3.1: Types of Toxic Effects
    - 2.6.6.3.2: Factors Affecting Toxicity
    - 2.6.6.3.3: Systemic Toxic Effects
    - 2.6.6.3.4: Organ Specific Toxic Effects
  - 2.6.6.4: Interactions
    - 2.6.6.4.1: Interactions
  - 2.6.6.5: Toxicity Testing Methods
    - 2.6.6.5.1: Testing and Assessing Toxicity
    - 2.6.6.5.2: Clinical Investigations and Other Types of Human Data
    - 2.6.6.5.3: Epidemiology Studies
  - 2.6.6.6: Risk Assessment
    - 2.6.6.6.1: Risk Assessment
    - 2.6.6.6.2: Hazard Identification
    - 2.6.6.6.3: Dose-Response Assessment
    - 2.6.6.6.4: Exposure Assessment
    - 2.6.6.6.5: Risk Characterization
  - 2.6.6.7: Exposure Standards and Guidelines
    - 2.6.6.7.1: Exposure Standards and Guidelines
    - 2.6.6.7.2: Regulation of Consumer Products and Drug Safety
    - 2.6.6.7.3: Environmental Exposure Standards/Guidelines
    - 2.6.6.7.4: Occupational (Workplace) Exposure Standards/Guidelines/Approaches
  - 2.6.6.8: Basic Physiology
    - 2.6.6.8.1: Introduction to Basic Physiology
    - 2.6.6.8.2: Homeostasis
    - 2.6.6.8.3: Organs and Organ Systems
    - 2.6.6.8.4: Tissues
    - 2.6.6.8.5: Cells

- 2.6.6.8.6: Chemicals
- 2.6.6.9: Introduction to Toxicokinetics
  - 2.6.6.9.1: What is Toxicokinetics
- 2.6.6.10: Absorption
  - 2.6.6.10.1: Introduction to Absorption
  - 2.6.6.10.2: Gastrointestinal Tract
  - 2.6.6.10.3: Respiratory Tract
  - 2.6.6.10.4: Dermal Route
  - 2.6.6.10.5: Other Routes of Exposure
- 2.6.6.11: Distribution
  - 2.6.6.11.1: Introduction to Distribution
  - 2.6.6.11.2: Influence of Route of Exposure
  - 2.6.6.11.3: Disposition Models
  - 2.6.6.11.4: Structural Barriers to Distribution
  - 2.6.6.11.5: Storage Sites
- 2.6.6.12: Biotransformation
  - 2.6.6.12.1: Introduction to Biotransformation
  - 2.6.6.12.2: Chemical Reactions
  - 2.6.6.12.3: Biotransformation Sites
  - 2.6.6.12.4: Modifiers of Biotransformation
- 2.6.6.13: Excretion
  - 2.6.6.13.1: Introduction to Secretion
  - 2.6.6.13.2: Urinary Excretion
  - 2.6.6.13.3: Fecal Excretion
  - 2.6.6.13.4: Exhaled Air
  - 2.6.6.13.5: Other Routes
- 2.6.6.14: Cellular Toxicology
  - 2.6.6.14.1: Adaptation
  - 2.6.6.14.2: Cell Damage and Tissue Repair
  - 2.6.6.14.3: Cancer
  - 2.6.6.14.4: Neurotoxicity
- 2.6.6.15: Intuitive Toxicology and Risk Communication
  - 2.6.6.15.1: Intuitive Toxicology
  - 2.6.6.15.2: Risk Communication
- 2.6.6.16: Environmental Toxicology, Environmental Health, and One Health
  - 2.6.6.16.1: Environmental Toxicology
  - 2.6.6.16.2: Environmental Health
  - 2.6.6.16.3: One Health
- 2.6.6.17: Conclusion
- Detailed Licensing

- Index
- Glossary
- Detailed Licensing

3: Ecology

- 3.1: Ecology Overview
- 3.2: Populations
  - 3.2.1: Population Dispersion
  - 3.2.2: Population Size
  - 3.2.3: Population Growth and Regulation
  - 3.2.4: Life History
  - 3.2.5: Data Dive- Wildlife Corridors
  - 3.2.6: Review
- 3.3: Communities
  - 3.3.1: Biotic Interactions
    - 3.3.1.1: Trophic Interactions
      - 3.3.1.1.1: Predation
      - 3.3.1.1.2: Herbivory
      - 3.3.1.1.3: Parasitism
      - 3.3.1.1.4: Food Chains and Food Webs
    - 3.3.1.2: Competition
    - 3.3.1.3: Facilitation
    - 3.3.1.4: Symbiosis
  - 3.3.2: Community Structure and Dynamics
  - 3.3.3: Data Dive- Bison Impacts In Prairies
  - 3.3.4: Review
- 3.4: Ecosystems
  - 3.4.1: Ecosystem Types and Dynamics
  - 3.4.2: Matter
  - 3.4.3: Biogeochemical Cycles
  - 3.4.4: Soils
  - 3.4.5: Soil Degradation
  - 3.4.6: Data Dive- Beaver Impacts on Wetlands
  - 3.4.7: Review
- 3.5: Biomes
  - 3.5.1: Climate and Biomes
  - 3.5.2: Terrestrial Biomes
  - 3.5.3: Aquatic Biomes
  - 3.5.4: Data Dive- Biome Carbon Storage
  - 3.5.5: Review

4: Conservation

- 4.1: The Value of Biodiversity
  - 4.1.1: Ecosystem Diversity
  - 4.1.2: Species Diversity
  - 4.1.3: Genetic Diversity
  - 4.1.4: Patterns of Biodiversity
  - 4.1.5: Data Dive- Biodiversity and Drugs
  - 4.1.6: Review
- 4.2: Threats to Biodiversity

- 4.2.1: Extinction
- 4.2.2: Measures of Biodiversity Loss
- 4.2.3: Habitat Loss
- 4.2.4: Overexploitation
- 4.2.5: Pollution
- 4.2.6: Invasive Species
- 4.2.7: Climate Change
- 4.2.8: Data Dive- Threats to Biodiversity
- 4.2.9: Review

- 4.3: Protecting Biodiveristy
  - 4.3.1: Policies
  - 4.3.2: Non-profit Organizations
  - 4.3.3: Species-level Conservation
  - 4.3.4: Protected Areas
  - 4.3.5: Ecosystem Restoration
  - 4.3.6: Economic Influences on Conservation
  - 4.3.7: Individual Choices
  - 4.3.8: Data Dive- Island Fox Populations
  - 4.3.9: Review

5: Humans and the Environment

- 5.1: The Human Population
  - 5.1.1: History of Human Population Growth
  - 5.1.2: The Rate of Human Population Growth
  - 5.1.3: Looking Ahead
  - 5.1.4: Data Dive- World Population Densities
  - 5.1.5: Review

- 5.2: Water Resources
  - 5.2.1: Fresh Water Supply and the Water Cycle
  - 5.2.2: Water Usage
  - 5.2.3: Water Scarcity and Solutions
  - 5.2.4: Data Dive- Aqueducts of Rome
  - 5.2.5: Review

- 5.3: Agriculture
  - 5.3.1: Food Security
  - 5.3.2: Industrial Agriculture
  - 5.3.3: Selective Breeding and Genetic Engineering
  - 5.3.4: Sustainable Agriculture
  - 5.3.5: Data Dive- Landslides and Ag Terraces
  - 5.3.6: Review

- 5.4: Environmental Health
  - 5.4.1: Types of Environmental Hazards
  - 5.4.2: Epidemiology
  - 5.4.3: Infectious Diseases
  - 5.4.4: Environmental Toxicology
  - 5.4.5: Environmental Hazard Reduction
  - 5.4.6: Data Dive- Cholera Cases Worldwide
  - 5.4.7: Review

6: Energy

- 6.1: Fossil Fuels
  - 6.1.1: Types of Fossil Fuels and Formation
  - 6.1.2: Mining, Processing, and Generating Electricity
  - 6.1.3: Fossil Fuel Consumption
  - 6.1.4: Consequences of Fossil Fuels
  - 6.1.5: Data Dive- Global Fossil Fuel Consumption
  - 6.1.6: Review
- 6.2: Nuclear Energy
  - 6.2.1: Radioactive Isotopes
  - 6.2.2: Generating Electricity with Nuclear Energy
  - 6.2.3: Nuclear Energy Consumption
  - 6.2.4: Consequences of Nuclear Energy
  - 6.2.5: Data Dive- Global Nuclear Power Generation
  - 6.2.6: Review
- 6.3: Renewable Energy
  - 6.3.1: Renewable Energy History and Consumption
  - 6.3.2: Wind Energy
  - 6.3.3: Solar Energy
  - 6.3.4: Geothermal Energy
  - 6.3.5: Hydropower
  - 6.3.6: Biofuels (Biomass Energy)
  - 6.3.7: Energy Conservation
  - 6.3.8: Data Dive- Global Renewable Energy Generation
  - 6.3.9: Review

7: Environmental Impacts

- 7.1: Solid Waste Management
  - 7.1.1: Waste Generation
  - 7.1.2: Waste Disposal
  - 7.1.3: Solid Waste and Marine Life
  - 7.1.4: Waste Reduction
  - 7.1.5: Data Dive- "Flushable" Wipes
  - 7.1.6: Review
- 7.2: Pollution
  - 7.2.1: Water Pollution
    - 7.2.1.1: Water Pollutants and Their Sources
    - 7.2.1.2: Water Treatment
    - 7.2.1.3: Mitigating Water Pollution
  - 7.2.2: Air Pollution
    - 7.2.2.1: Ozone Depletion
    - 7.2.2.2: Acid Deposition
  - 7.2.3: Data Dive- Plastic in World Oceans
  - 7.2.4: Review
- 7.3: Climate Change
  - 7.3.1: The Greenhouse Effect and Climate Change

- 7.3.2: Data Dive- Flooding in Venice Italy
- 7.3.3: Review

8: Sustainability

- 8.1: Introduction to Sustainability
- 8.2: History of Sustainability
- 8.3: Environmental Ethics
  - 8.3.1: Environmental Justice and Indigenous Struggles
  - 8.3.2: Perspectives of the Environment
- 8.4: Environmental Policy
- 8.5: Sustainable Cities
- 8.6: Sustainability and the Future
- 8.7: Data Dive- Sustainable Investments
- 8.8: Review

Index

Glossary

Detailed Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

CHAPTER OVERVIEW

1: Introduction

The objective of this unit is to not only get you acquainted with what this book will cover but to also introduce the rigors and structure of science and how mathematics plays a crucial role in helping scientists interpret results in the most unbiased way possible. Science and mathematics are not just for scientists. A base understanding of these subjects is crucial in our very complex modern world. Without a solid understanding of how science and mathematics play a role in our everyday lives, we cannot respond appropriately in everyday situations, know how to vote on scientific laws, or hold public representatives responsible for their decisions on scientific policy.

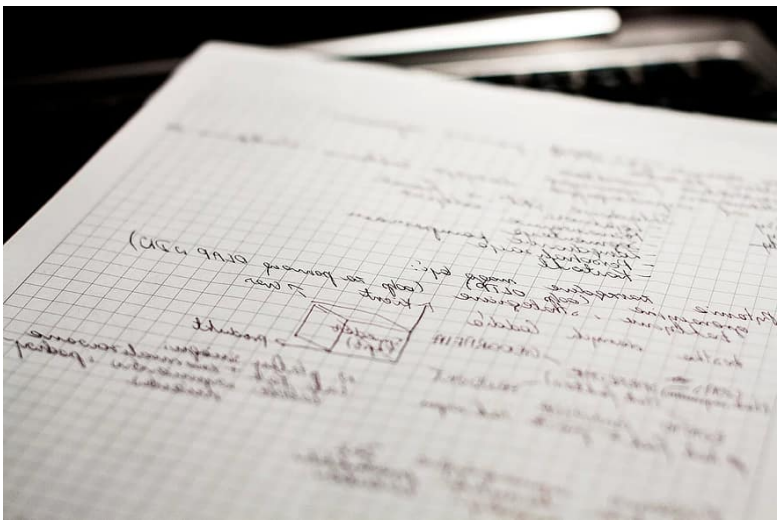


Figure $1.a$ Written notes. Image by Pikist (Public domain)

Attributions

Rachel Schleiger (CC-BY-NC)

1.1: Environmental Science Preface

1.2: What Is Science, and How Does It Work?

1.2.1: What is Science?

1.2.2: The Scientific Method

1.2.3: Scientific Papers

1.2.4: Basic and Applied Science

1.2.5: Data Dive- Tonsillectomy Trends

1.2.6: Review

1.3: Math Blast- An Overview of Essential Mathematics Used in Science

1.3.1: Getting down the basics- What are the essential mathematical and statistical terms and concepts?

1.3.2: Summarizing the data- Descriptive statistics

1.3.3: Putting it all together- Inferential statistics and hypothesis testing

1.3.4: Data Dive- California Wildfires

1.3.5: Review

This page titled 1: Introduction is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

1.1: Environmental Science Preface

What is Environmental Science?

Environmental science is the dynamic, interdisciplinary study of the interaction of living and non-living parts of the environment, with special focus on the impact of humans on the environment. The study of environmental science includes circumstances, objects, or conditions by which an organism or community is surrounded and the complex ways in which they interact.

Environmental Science is Interdisciplinary

Environmental science includes disciplines in the physical sciences (like geology, soil science, physical geography, chemistry, and atmospheric sciences), life sciences (like ecology, conservation biology, restoration biology, and population biology), social sciences (like human geography, economics, law, political science, and anthropology), and humanities (philosophy, ethics). As such, environmental scientists are a diverse bunch. However, they all come together for the focus of studying and identifying past, current, and future environmental issues while also exploring solutions to the environmental issues and also considering practical needs.

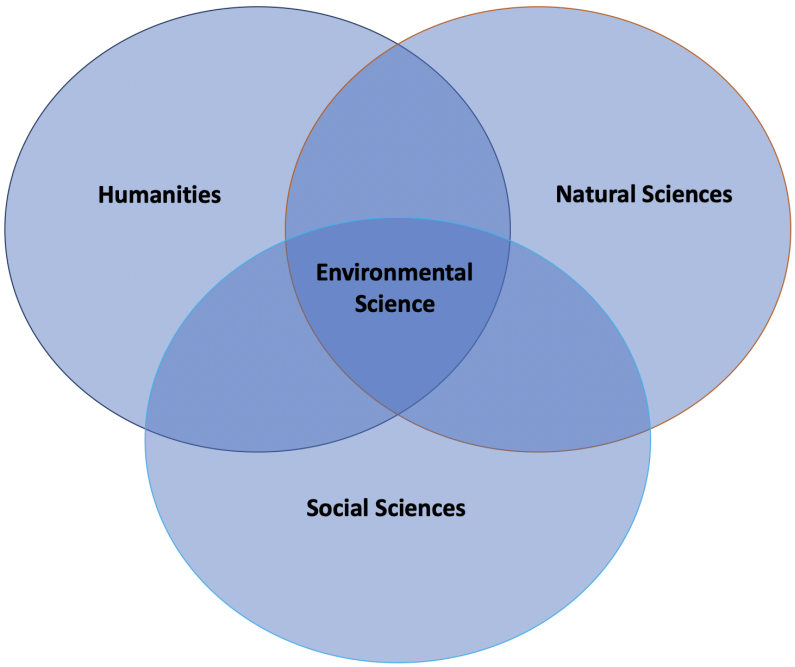


Figure 1.1.*a*: Venn diagram illustrating how environmental science is interdisciplinary between humanities, natural sciences (which includes physical and life sciences) and social sciences. Image by Rachel Schleiger (CC-BY-NC).

Why Study Environmental Science?

The need for equitable, ethical, and sustainable use of Earth's resources by a global population that nears the carrying capacity of the planet requires us not only to understand how human behaviors affect the environment, but also the scientific principles that govern interactions between the living and non-living. Our future depends on our ability to understand and evaluate evidence-based arguments about the environmental consequences of human actions and technologies, and to make informed decisions based on those arguments.

From global climate change to habitat loss driven by human population growth and development, Earth is becoming a different planet—right before our eyes. The global scale and rate of environmental change are beyond anything in recorded human history. Our challenge is to acquire an improved understanding of Earth's complex environmental systems; systems characterized by

interactions within and among their natural and human components that link local to global and short-term to long-term phenomena, and individual behavior to collective action. The complexity of environmental challenges demands that we all participate in finding and implementing solutions leading to long-term environmental sustainability.

The Tragedy of the Commons

In his essay, The Tragedy of the Commons, Garrett Hardin (1968) looked at what happens when humans do not limit their actions by including the land as part of their ethic. The **tragedy of the commons** develops in the following way: Imagine a pasture open to all (the 'commons'). It is to be expected that each herder will try to keep as many cattle as possible on the commons. As rational beings, each herdsman seeks to maximize their gain. Adding more cattle increases their profit, and they do not suffer any immediate negative consequence because the commons are shared by all. The rational herder concludes that the only sensible course is to add another animal to their herd, and then another, and so forth. However, this same conclusion is reached by each and every rational herder sharing the commons. Therein lies the tragedy: each person is locked into a system that compels them to increase their herd, without limit, in a world that is limited. Eventually this leads to the degradation of the commons. In a society that believes in the freedom of the commons, freedom brings ruin to all because each person acts selfishly.

Hardin went on to apply the situation to modern commons: overgrazing of public lands, overuse of public forests and parks, depletion of fish populations in the ocean, use of rivers as a common dumping ground for sewage, and releasing air pollutants.

The "Tragedy of the Commons" is applicable to what is arguably the most consequential environmental problem: global climate change. The atmosphere is a commons into which countries are dumping carbon dioxide from the burning of fossil fuels. Although we know that the generation of greenhouse gases will have damaging effects upon the entire globe, we continue to burn fossil fuels. As a country, the immediate benefit from the continued use of fossil fuels is seen as a positive component (because of economic growth). All countries, however, will share the negative long-term effects.

Some Indicators of Global Environmental Stress

- **Forests** — Deforestation remains a main issue. 1 million hectares of forest were lost every year in the decade 1980-1990. The largest losses of forest area are taking place in the tropical moist deciduous forests, the zone best suited to human settlement and agriculture. Recent estimates suggest that nearly two-thirds of tropical deforestation is due to farmers clearing land for agriculture. There is increasing concern about the decline in forest quality associated with intensive use of forests and unregulated access.
- **Soil** — As much as 10% of the earth's vegetated surface is now at least moderately degraded. Trends in soil quality and management of irrigated land raise serious questions about longer-term sustainability. It is estimated that about 20% of the world's 250 million hectares of irrigated land are already degraded to the point where crop production is seriously reduced.
- **Fresh water** — Some 20% of the world's population lacks access to safe water and 50% lacks access to safe sanitation. If current trends in water use persist, two-thirds of the world's population could be living in countries experiencing moderate or high water stress by 2025.
- **Marine fisheries** — 25% of the world's marine fisheries are being fished at their maximum level of productivity and 35% are overfished (yields are declining). In order to maintain current per capita consumption of fish, global fish harvests must be increased; much of the increase might come through aquaculture which is a known source of water pollution, wetland loss and mangrove swamp destruction.
- **Biodiversity** — Biodiversity is increasingly coming under threat from development, which destroys or degrades natural habitats, and from pollution from a variety of sources. The first comprehensive global assessment of biodiversity put the total number of species at close to 14 million and found that between 1% and 11% of the world's species may be threatened by extinction every decade. Coastal ecosystems, which host a very large proportion of marine species, are at great risk with perhaps one-third of the world's coasts at high potential risk of degradation and another 17% at moderate risk.
- **Atmosphere** — The Intergovernmental Panel on Climate Change has established that human activities are having a discernible influence on global climate. CO_2 emissions in most industrialized countries have risen during the past few years and countries generally failed to stabilize their greenhouse gas emissions at 1990 levels by 2000 as required by the Climate Change convention.

- **Toxic chemicals** — About 100,000 chemicals are now in commercial use and their potential impacts on human health and ecological function represent largely unknown risks. Persistent organic pollutants are now so widely distributed by air and ocean currents that they are found in the tissues of people and wildlife everywhere; they are of particular concern because of their high levels of toxicity and persistence in the environment.
- **Hazardous wastes** — Pollution from heavy metals, especially from their use in industry and mining, is also creating serious health consequences in many parts of the world. Incidents and accidents involving uncontrolled radioactive sources continue to increase, and particular risks are posed by the legacy of contaminated areas left from military activities involving nuclear materials.
- **Waste** — Domestic and industrial waste production continues to increase in both absolute and per capita terms, worldwide. In the developed world, per capita waste generation has increased threefold over the past 20 years; in developing countries, it is highly likely that waste generation will double during the next decade. The level of awareness regarding the health and environmental impacts of inadequate waste disposal remains rather poor; poor sanitation and waste management infrastructure is still one of the principal causes of death and disability for the urban poor.

Attributions

Modified by Melissa Ha and Rachel Schleiger from The Earth, Humans, & the Environment, Environment and Sustainability, and Environmental Ethics from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 1.1: Environmental Science Preface is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **1: Environmental Science Preface** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

1.2: What Is Science, and How Does It Work?

Chapter Hook

It's 1935, and you have another sore throat, the second time this month. After a doctor appointment you are referred to an otolaryngologist (ear, nose, and throat doctor). After a short appointment you are scheduled for a tonsillectomy the next day. The majority of children during this time period had their tonsils removed. In fact, the history of tonsillectomies goes back 2000 years! Medical professionals knew that taking out tonsils had benefits for their patients, mainly decreased throat infections and improved sleep due to obstructive sleep apnea. However, there was not extensive research performed on long term risks in addition to the surgical risks. By the 1970s tonsillectomy recommendations significantly decreased after research confirmed long term side effects for some patients. This is one of many examples where procedures and policies were created without appropriate evidence utilizing the scientific method, the backbone of science.

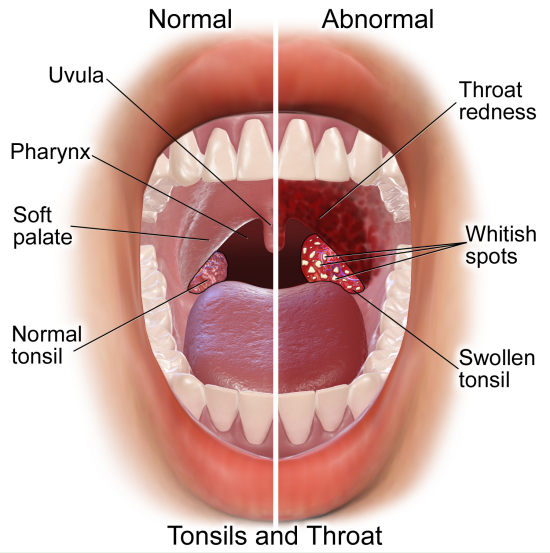


Figure 1.2.*a* Anatomy of normal versus abnormal tonsils and throat. Image by BruceBlaus (licensed under CC-BY-3.0)

Attribution

Chapter Hook by Rachel Schleiger (CC-BY-NC)

Page summary modified by Melissa Ha from What is Science? from *An Introduction to Geology* by Chris Johnson et al. (licensed under CC-BY-NC-SA)

1.2.1: What is Science?

1.2.2: The Scientific Method

1.2.3: Scientific Papers

1.2.4: Basic and Applied Science

1.2.5: Data Dive- Tonsillectomy Trends

1.2.6: Review

This page titled 1.2: What Is Science, and How Does It Work? is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

1.2.1: What is Science?

Like other natural sciences, environmental science gathers knowledge about the natural world. Science is more than just a body of knowledge, science provides a means to evaluate and create new knowledge. The methods of science include careful observation, record keeping, logical and mathematical reasoning, experimentation, and submitting conclusions to the scrutiny of others. Science also requires considerable imagination and creativity; a well-designed experiment is commonly described as elegant or beautiful. Science has considerable practical implications and some science is dedicated to practical applications, such as the prevention of disease (figure 1.2.1.*a*). Other science proceeds largely motivated by curiosity. Whatever its goal, there is no doubt that science has transformed human existence and will continue to do so.

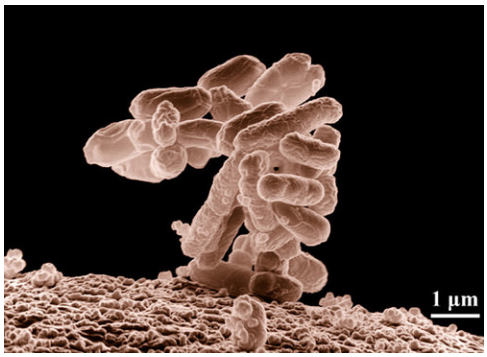


Figure 1.2.1.*a*: Biologists may choose to study *Escherichia coli* (*E. coli*), a bacterium that is a normal resident of our digestive tracts but which is also sometimes responsible for disease outbreaks. In this micrograph, the bacterium is visualized using a scanning electron microscope and digital colorization. (credit: Eric Erbe; digital colorization by Christopher Pooley, USDA-ARS)

There are areas of knowledge, however, to which the methods of science cannot be applied. These include such things as morality, aesthetics, or spirituality. Science cannot investigate these areas because they are outside the realm of material phenomena, the phenomena of matter and energy, and cannot be observed and measured.

Evidence, Measurements, and Observations

Scientists use objective evidence over subjective evidence, to reach sound and logical conclusions. An **objective** observation is without personal bias and the same by all individuals. **Bias** refers to favoring one thing over another, and it can lead to inaccurate results. Humans are biased by nature, so they cannot be completely objective; the goal is to be as unbiased as possible. A **subjective** observation is based on a person's feelings and beliefs and is unique to that individual (figure 1.2.1.*b*).



Figure 1.2.1.*b*: This is Grand Canyon of the Yellowstone in Yellowstone National Park. An objective statement about this would be, "The picture is of a waterfall." A subjective statement would be, "The picture is beautiful."

Another way scientists avoid bias is by using quantitative over qualitative measurements whenever possible. A **quantitative measurement** is expressed with a specific numerical value. **Qualitative observations** are general or relative descriptions. For example, describing a rock as red or heavy is a qualitative observation. Determining a rock's color by measuring wavelengths of reflected light or its density by measuring the proportions of minerals it contains is quantitative. Numerical values are more precise than general descriptions, and they can be analyzed using statistical calculations. This is why quantitative measurements are much more useful to scientists than qualitative observations.

Inductive and Deductive Reasoning

One thing is common to all forms of science: an ultimate goal to know. Curiosity and inquiry are the driving forces for the development of science. Scientists seek to understand the world and the way it operates. Two methods of logical thinking are used: inductive reasoning and deductive reasoning.

Inductive reasoning is a form of logical thinking that uses related observations to arrive at a general conclusion. This type of reasoning is common in **descriptive science**. A life scientist such as a biologist makes observations and records them. The raw data can be supplemented with drawings, pictures, photos, or videos. From many observations, the scientist can infer conclusions (inductions) based on evidence. Inductive reasoning involves formulating generalizations inferred from careful observation and the analysis of a large amount of data. Surveying land use (which areas are forested, agricultural, urban, etc.) across the United States and then concluding that forested areas are concentrated in the West is an example of descriptive science.

In deductive reasoning, the pattern of thinking moves in the opposite direction as compared to inductive reasoning. **Deductive reasoning** is a form of logical thinking that uses a general principle or law to forecast specific results. From those general principles, a scientist can extrapolate and predict the specific results that would be valid as long as the general principles are valid. For example, a prediction would be that if the climate is becoming warmer in a region, the distribution of plants and animals should change. Comparisons have been made between distributions in the past and the present, and the many changes that have been found are consistent with a warming climate. Finding the change in distribution is evidence that the climate change conclusion is a valid one. Deductive reasoning, or deduction, is the type of logic used in hypothesis-based science (see below).

To summarize, inductive reasoning moves from the specific (an observation) to general (conclusion), and deductive reasoning moves from the general (a hypothesis or principle) to the specific (results).

Both types of logical thinking are related to the two main pathways of scientific study: descriptive science and hypothesis-based science. **Descriptive** (or discovery) **science** aims to observe, explore, and discover, while **hypothesis-based science** begins with a specific question or problem and a potential answer or solution that can be tested. The boundary between these two forms of study is often blurred, because most scientific endeavors combine both approaches. Observations lead to questions, questions lead to forming a hypothesis as a possible answer to those questions, and then the hypothesis is tested. Thus, descriptive science and hypothesis-based science are in continuous dialogue.

Science is also a Social Process

Scientists share their ideas with peers at conferences, seeking guidance and feedback (figure 1.2.1.*c*). Research papers and data submitted for publication are rigorously reviewed by qualified peers, scientists who are experts in the same field. The scientific review process aims to weed out misinformation, invalid research results, and wild speculation. Thus, it is slow, cautious, and conservative. Scientists tend to wait until a hypothesis is supported by an overwhelming amount of evidence from many independent researchers before accepting it as a **scientific theory**.



Figure 1.2.1.*c*: Scientists share information by publishing and attending conferences. The conference shown here focuses on peanut and mycotoxin (harmful chemicals produced by fungi) research. Image by Sharon Dowdy (CC-BY-NC).

Characteristics of Scientists

There is nothing mysterious or even particularly unusual about the things that scientists do. There are many ways to work on scientific problems. They all require common sense. Beyond that, they all display certain features that are especially - but not uniquely - characteristic of science.

- Skepticism — Good scientists use highly-critical standards in the judging of evidence. They approach data, claims, and theories (ideally, even their own!) with healthy doses of skepticism.
- Tolerance of uncertainty — Scientists often work for years - sometimes for an entire career - trying to understand one scientific problem. This often involves finding facts that, for a time, fail to fit into any coherent pattern and that even may support mutually contradictory explanations. Sometimes, as one listens to scientists vigorously defending their views, their confidence seems absolute. But deep in their hearts, they know that their views are based on probabilities and that a new piece of evidence may turn up at any time and force a major shift in their views.
- Although they certainly have no monopoly on hard work, their willingness to work long hours and years pursuing a problem is the mark of all good scientists. For science is hard work.

Attribution

Modified by Melissa Ha from the following sources:

- What is Science? from *An Introduction to Geology* by Chris Johnson et al. (licensed under CC-BY-NC-SA)
- The Process of Science from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Scientific Methods from *Biology* by John W. Kimball (licensed under CC-BY)

This page titled 1.2.1: What is Science? is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **2.1: What is Science?** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

1.2.2: The Scientific Method

The **scientific method** is a process of research with defined steps that include data collection and careful observation.

Observation

Scientific advances begin with **observations**. This involves noticing a pattern, either directly or indirectly from the literature. An example of a direct observation is noticing that there have been a lot of toads in your yard ever since you turned on the sprinklers, where as an indirect observation would be reading a scientific study reporting high densities of toads in urban areas with watered lawns.

During the Vietnam War (figure $1.2.2.a$), press reports from North Vietnam documented an increasing rate of birth defects. While this credibility of this information was initially questioned by the U.S., it evoked questions about what could be causing these birth defects. Furthermore, increased incidence of certain cancers and other diseases later emerged in Vietnam veterans who had returned to the U.S. This leads us to the next step of the scientific method, the question.

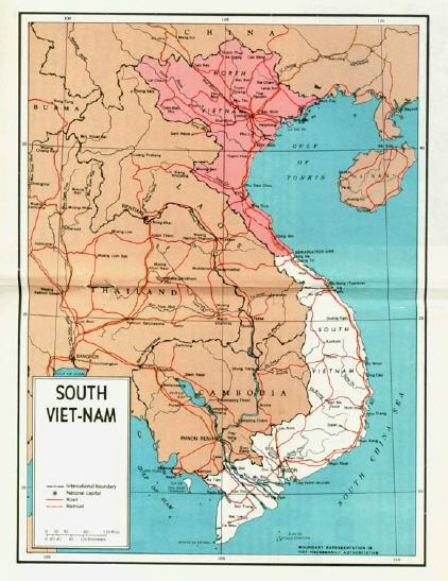


Figure $1.2.2.a$: A map of Vietnam 1954-1975. Image from Bureau of Public Affairs U.S. Government Printing Office (public domain).

Question

The **question** step of the scientific method is simply asking, what explains the observed pattern? Multiple questions can stem from a single observation. Scientists and the public began to ask, what is causing the birth defects in Vietnam and diseases in Vietnam veterans? Could it be associated with the widespread military use of the herbicide Agent Orange to clear the forests (figure $1.2.2.b-c$), which helped identify enemies more easily?



Figure $1.2.2.b$: Agent Orange drums in Vietnam. Image by U.S. Government (public domain).



Figure 1.2.2.*c*: A healthy mangrove forest (top), and another forest after application of Agent Orange. Image by unknown author (public domain).

Hypothesis and Prediction

The **hypothesis** is the expected answer to the question. The best hypotheses state the proposed direction of the effect (increases, decreases, etc.) and explain why the hypothesis could be true.

- OK hypothesis: Agent Orange influences rates of birth defects and disease.
- Better hypothesis: Agent Orange increases the incidence of birth defects and disease.
- Best hypothesis: Agent Orange increases the incidence of birth defects and disease because these health problems have been frequently reported by individuals exposed to this herbicide.

If two or more hypotheses meet this standard, the simpler one is preferred.

Predictions stem from the hypothesis. The prediction explains what results would support hypothesis. The prediction is more specific than the hypothesis because it references the details of the experiment. For example, "If Agent Orange causes health problems, then mice experimentally exposed to TCDD, a contaminant of Agent Orange, during development will have more frequent birth defects than control mice" (figure 1.2.2.*d*).

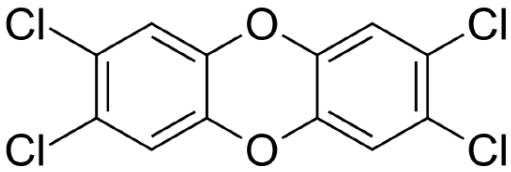


Figure 1.2.2.*d*: The chemical structure of TCDD (2,3,7,8-tetrachlorodibenzo-p-dioxin), which is produced when synthesizing the chemicals in Agent Orange. It contaminates Agent Orange at low but harmful concentrations. Image by Emeldir (public domain).

Hypotheses and predictions must be testable to ensure that it is valid. For example, a hypothesis that depends on what a bear thinks is not testable, because it can never be known what a bear thinks. It should also be **falsifiable**, meaning that they have the capacity to be tested and demonstrated to be untrue. An example of an unfalsifiable hypothesis is “Botticelli’s *Birth of Venus* is beautiful.” There is no experiment that might show this statement to be false. To test a hypothesis, a researcher will conduct one or more experiments designed to eliminate one or more of the hypotheses. This is important. A hypothesis can be disproven, or eliminated, but it can never be proven. Science does not deal in proofs like mathematics. If an experiment fails to disprove a hypothesis, then we find support for that explanation, but this is not to say that down the road a better explanation will not be found, or a more carefully designed experiment will be found to falsify the hypothesis.

Hypotheses are tentative explanations and are different from scientific theories. A **scientific theory** is a widely-accepted, thoroughly tested, and confirmed explanation for a set of observations or phenomena. Scientific theory is the foundation of scientific knowledge. In addition, in many scientific disciplines (less so in biology) there are **scientific laws**, often expressed in mathematical formulas, which describe how elements of nature will behave under certain specific conditions, but they do not offer explanations for why they occur.

Design an Experiment

Next, a scientific study (experiment) is planned to test the hypothesis and determine whether the results match the predictions. Each experiment will have one or more variables. The **independent variable** is what scientists hypothesize might be causing something else. In a manipulative experiment (see below), the independent variable is manipulated by the scientist. The **dependent variable** is the response, the variable ultimately measured in the study. **Controlled variables** (confounding factors) might affect the dependent variable, but they are not the focus of the study. Scientist attempt to standardize the controlled variables so that they do not influence the results. In our previous example, exposure to Agent Orange is the independent variable. It is hypothesized to cause a change in health (likelihood of having children with birth defects or developing a disease), the dependent variable. Many other things could affect health, including diet, exercise, and family history. These are the controlled variables.

There are two main types of scientific studies: experimental studies (manipulative experiments) and observational studies.

In a **manipulative experiment**, the independent variable is altered by the scientists, who then observe the response. In other words, the scientists apply a **treatment**. An example would be exposing developing mice to TCDD and comparing the rate of birth defects to a control group. The **control group** is group of test subjects that are as similar as possible to all other test subjects, with the exception that they don't receive the experimental treatment (those that do receive it are known as the experimental, treatment, or **test group**). The purpose of the control group is to establish what the dependent variable would be under normal conditions, in the absence of the experimental treatment. It serves as a baseline to which the test group can be compared. In this example, the control group would contain mice that were not exposed to TCDD but were otherwise handled the same way as the other mice (figure 1.2.2.e)



Figure 1.2.2.e: Laboratory mice. In a proper scientific study, the treatment would be applied to multiple mice. Another group of mice would not receive the treatment (the control group). Image by Aaron Logan (CC-BY).

In an **observational study**, scientists examine multiple samples with and without the presumed cause. An example would be monitoring the health of veterans who had varying levels of exposure to Agent Orange.

Scientific studies contain many **replicates.** Multiple samples ensure that any observed pattern is due to the treatment rather than naturally occurring differences between individuals. A scientific study should also be **repeatable**, meaning that if it is conducted again, following the same procedure, it should reproduce the same general results. Additionally, multiple studies will ultimately test the same hypothesis.

Results

Finally, the data are collected and the results are analyzed. As described in the Math Blast chapter, statistics can be used to describe the data and summarize data. They also provide a criterion for deciding whether the pattern in the data is strong enough to support

the hypothesis.

The manipulative experiment in our example found that mice exposed to high levels of 2,4,5-T (a component of Agent Orange) or TCDD (a contaminant found in Agent Orange) during development had a cleft palate birth defect more frequently than control mice (figure 1.2.2.f). Mice embryos were also more likely to die when exposed to TCDD compared to controls.



Figure 1.2.2.f: Cleft lip and palate, a birth defect in which these structures are split. Image by James Heilman, MD (CC-BY-SA).

An observational study found that self-reported exposure to Agent Orange was positively correlated with incidence of multiple diseases in Korean veterans of the Vietnam War, including various cancers, diseases of the cardiovascular and nervous systems, skin diseases, and psychological disorders. Note that a **positive correlation** simply means that the independent and dependent variables both increase or decrease together, but further data, such as the evidence provided by manipulative experiments is needed to document a **cause-and-effect relationship**. (A **negative correlation** occurs when one variable increases as the other decreases.)

Conclusion

Lastly, scientists make a conclusion regarding whether the data support the hypothesis. In the case of Agent Orange, the data, that mice exposed to TCDD and 2,4,5-T had higher frequencies of cleft palate, matches the prediction. Additionally, veterans exposed to Agent Orange had higher rates of certain diseases, further supporting the hypothesis. We can thus accept the hypothesis that Agent Orange increases the incidence of birth defects and disease.

In practice, the scientific method is not as rigid and structured as it might first appear. Sometimes an experiment leads to conclusions that favor a change in approach; often, an experiment brings entirely new scientific questions to the puzzle. Many times, science does not operate in a linear fashion; instead, scientists continually draw inferences and make generalizations, finding patterns as their research proceeds (figure 1.2.2.g). Even if the hypothesis was supported, scientists may still continue to test it in different ways. For example, scientists explore the impacts of Agent Orange, examining long-term health impacts as Vietnam veterans age.

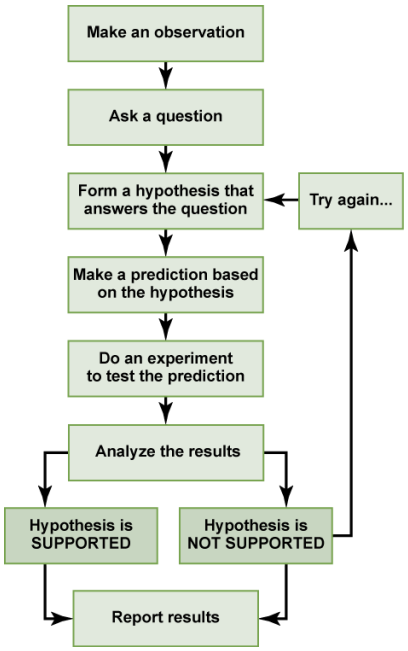


Figure 1.2.2.*g*: The scientific method is a series of defined steps that include experiments and careful observation. The steps are as follows: make an observation; ask a question; form a hypothesis that answers the question; make a prediction based on the hypothesis; do an experiment to test the prediction; analyze the results; and report the results. Whether the hypothesis is supported or not supported, the results are still reported. If a hypothesis is not supported by data, a new hypothesis can be proposed.

Scientific findings can influence decision making. In response to evidence regarding the effect of Agent Orange on human health, compensation is now available for Vietnam veterans who were exposed to Agent Orange and develop certain diseases. The use of Agent Orange is also banned in the U.S. Finally, the U.S. has began cleaning sites in Vietnam that are still contaminated with TCDD.

Building on the Work of Others

Only rarely does a scientific discovery spring full-blown on the scene. When it does, it is likely to create a revolution in the way scientists perceive the world around them and to open up new areas of scientific investigation. Darwin's theory of evolution and Mendel's rules of inheritance are examples of such revolutionary developments. Most science, however, consists of adding another brick to an edifice that has been slowly and painstakingly constructed by prior work.

The development of a new technique often lays the foundation for rapid advances along many different scientific avenues. Just consider the advances in biology that discovery of the light microscope and, later, the electron microscope have made possible. Throughout these pages, there are many examples of experimental procedures. Each was developed to solve a particular problem. However, each was then taken up by workers in other laboratories and applied to their problems.

In a similar way, the creation of a new explanation (hypothesis) in a scientific field often stimulates workers in related fields to reexamine their own field in the light of the new ideas. Darwin's theory of evolution, for example, has had an enormous impact on virtually every subspecialty in biology as well as environmental science. To this very day, scientists in specialties as different as biochemistry and conservation biology are guided in their work by evolutionary theory (figure 1.2.2.*g*).



Figure 1.2.2.h: Understanding the endangered mountain yellow-legged frog through the lens of evolution and genetics has aided its conservation. Image by Iasaac Chellman/NPS (CC-BY).

References

Institute of Medicine (US) Committee to Review the Health Effects in Vietnam Veterans of Exposure to Herbicides. Veterans and Agent Orange: Health Effects of Herbicides Used in Vietnam. Washington (DC): National Academies Press (US); 1994. 2, History of the Controversy Over the Use of Herbicides.

Neubert, D., Dillmann, I. Embryotoxic effects in mice treated with 2,4,5-trichlorophenoxyacetic acid and 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **272,** 243–264 (1972).

Stellman, J. M., & Stellman, S. D. (2018). Agent Orange During the Vietnam War: The Lingering Issue of Its Civilian and Military Health Impact. *American journal of public health*, *108*(6), 726–728.

Yi, S. W., Ohrr, H., Hong, J. S., & Yi, J. J. (2013). Agent Orange exposure and prevalence of self-reported diseases in Korean Vietnam veterans. *Journal of preventive medicine and public health = Yebang Uihakhoe chi*, *46*(5), 213–225.

Attributions

Modified by Melissa Ha from the following sources:

- The Process of Science from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Scientific Methods from *Biology* by John W. Kimball (licensed under CC-BY)

This page titled 1.2.2: The Scientific Method is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **2.2: The Scientific Method** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

1.2.3: Scientific Papers

This section discusses what peer review is and how scientific papers are structured.

Peer Review

Scientists must share their findings for other researchers to expand and build upon their discoveries. For this reason, an important aspect of a scientist's work is disseminating results and communicating with peers. Scientists can share results by presenting them at a scientific meeting or conference, but this approach can reach only the limited few who are present. Instead, most scientists present their results in peer-reviewed articles that are published in scientific journals (figure 1.2.3.a). **Peer-reviewed articles** are scientific papers that are evaluated, usually anonymously by a scientist's colleagues, or peers. These colleagues are qualified individuals, often experts in the same research area, who judge whether or not the scientist's work is suitable for publication. Peer reviewers assess experimental design, statistical analysis, presentation of data, and whether conclusions fit the results. The process of peer review helps to ensure that the research described in a scientific paper is original, significant, logical, ethical, and thorough. Scientists publish their work so other scientists can reproduce their experiments under similar or different conditions to expand on the findings. The experimental results must be consistent with the findings of other scientists.



Figure 1.2.3.a: The findings of scientific studies are published in peer-reviewed scientific journals. Image by free svg (public domain).

As you review scientific information, whether in an academic setting or as part of your day-to-day life, it is important to think about the credibility of that information. You might ask yourself: has this scientific information been through the rigorous process of peer review? Are the conclusions based on available data and accepted by the larger scientific community? Scientists are inherently skeptical, especially if conclusions are not supported by evidence (and you should be too).

Structure of Scientific Papers

Scientific papers are usually divided into several sections (not necessarily in this order).

Summary or Abstract

This section includes only the essence of the other sections. It should be as brief as possible, telling the reader what the goal of the experiment was, what was found, and the significance of the findings. The abstract is often placed at the beginning of the paper rather than at its end.

Introduction

This section of the paper describes the scientific question or problem that was the subject of the investigation. The introduction also includes references to earlier reports of these and other scientists that have served as the foundation for the present work. Finally, the introduction states the hypothesis.

Materials and Methods

Here are precisely described the materials used (e.g., strains of organism, source of the reagents) and all the methods followed. The goal of this section is to give all the details necessary for workers in other laboratories to be able to **repeat the experiments exactly**. When many complex procedures are involved, it is acceptable to refer to earlier papers describing these methods in greater detail.

Results

Here the authors report what happened in their experiments. This report is usually supplemented with graphs, tables, and photographs.

Discussion

Here the authors point out what they think is the significance of their findings. This is the place to show that the results are compatible with certain hypotheses and less compatible, or even incompatible, with others. If the results contradict the results of similar experiments in other laboratories, the discrepancies are noted here, and an attempt may be made to reconcile the differences.

Acknowledgments

In this brief but important section, the authors give credit to those who have assisted them in the work. These usually include technicians (who may have actually performed most of the experiments!) and other scientists who donated materials for the experiments and/or gave advice about them.

References

This section gives a careful listing of all earlier scientific work referred to in the main body of the paper. Most of the references are to other scientific papers. Each reference should provide enough information so that another person can locate the document. This means that each reference should include the name(s) of the author(s), the journal or book in which the report appears, and the year of publication. In the case of scientific journals, the volume number in which the paper appears and the page number on which the paper begins should be included. A digital object identifier (DOI) or url is often included. The full title of the article is often included as well, although some citation styles omit the title from the reference.

Attribution

Modified by Melissa Ha from the following sources:

- Scientific Papers from *Biology* by John W. Kimball (CC-BY)
- The Process of Science from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 1.2.3: Scientific Papers is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **2.3: Scientific Papers** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

1.2.4: Basic and Applied Science

Is it valuable to pursue science for the sake of simply gaining knowledge, or does scientific knowledge only have worth if we can apply it to solving a specific problem or bettering our lives? This question focuses on the differences between two types of science: basic science and applied science.

Basic science or “pure” science seeks to expand knowledge regardless of the short-term application of that knowledge. It is not focused on developing a product or a service of immediate public or commercial value. The immediate goal of basic science is knowledge for knowledge’s sake, though this does not mean that in the end it may not result in an application. Questions such as, "How have plants evolved to attract pollinators?" and "What factors determine which species will co-occur with each other?" fall under the scope of basic science (figure 1.2.4.*a*).



Figure 1.2.4.*a*: Many different plant species are co-occurring (growing together) at the Brisbane Botanic Gardens. Exploring which plants co-occur naturally falls under the scope of basic science. Image by Brisbane City Council (CC-BY).

In contrast, **applied science** aims to use science to solve real-world problems, such as improving crop yield, find a cure for a particular disease, or save animals threatened by a natural disaster. In applied science, the problem is usually defined for the researcher.

A layperson (nonscientist) may perceive applied science as “useful” and basic science as “useless”, posing the question, "What for?" to a scientist advocating knowledge acquisition. A careful look at the history of science, however, reveals that basic knowledge has resulted in many remarkable applications of great value. Many scientists think that a basic understanding of science is necessary before an application is developed; therefore, applied science relies on the results generated through basic science. Other scientists think that it is time to move on from basic science and instead to find solutions to actual problems. Both approaches are valid. It is true that there are problems that demand immediate attention; however, few solutions would be found without the help of the knowledge generated through basic science.

One example of how basic and applied science can work together to solve practical problems occurred after the discovery of DNA structure led to an understanding of the molecular mechanisms governing DNA replication. Strands of DNA, unique in every human, are found in our cells, where they provide the instructions necessary for life. During DNA replication, new copies of DNA are made, shortly before a cell divides to form new cells. Understanding the mechanisms of DNA replication (through basic science) enabled scientists to develop laboratory techniques that are now used to identify genetic diseases, pinpoint individuals who were at a crime scene, and determine paternity (all examples of applied science). Without basic science, it is unlikely that applied science would exist.

Another example of the link between basic and applied research is the Human Genome Project, a study in which each human chromosome was analyzed and mapped to determine the precise sequence of the DNA code and the exact location of each gene. (The gene is the basic unit of heredity; an individual’s complete collection of genes is his or her genome.) Other organisms have also been studied as part of this project to gain a better understanding of human chromosomes. The Human Genome Project (figure 1.2.4.*b*) relied on basic research carried out with non-human organisms and, later, with the human genome. An important end goal eventually became using the data for applied research seeking cures for genetic diseases.

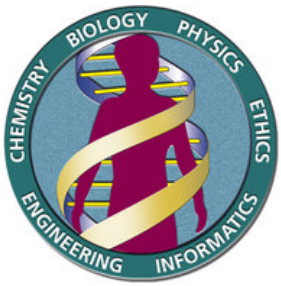


Figure 1.2.4.*b*: The Human Genome Project was a 13-year collaborative effort among researchers working in several different fields of science. The project was completed in 2003. (credit: the U.S. Department of Energy Genome Programs)

The discovery of penicillin, the first antibiotic, also originated in basic science. The mold *Penicillium* (figure 1.2.4.*c*) contaminated petri dishes of bacteria and unexpectedly inhibited their growth. Read more about The Real Story behind Penicillin.



Figure 1.2.4.*c*: The mold *Penicillium* growing on a petri dish. Image by Crulina 98 (CC-BY-SA).

Ecological modeling, which is closely intertwined with the field of environmental science, is another example in which applied science closely relies on basic science. **Ecological models** are complex equations through which computers can predict the outcome of different decisions or scenarios based on existing data. For example, a forest manager might use a model to determine which pattern of tree removal will promote forest health and produce a steady, sustainable supply of timber. The data used to create the ecological model are collected through a combination of basic and applied studies.

Attribution

Modified by Melissa Ha from The Process of Science from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 1.2.4: Basic and Applied Science is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **2.4: Basic and Applied Science** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

1.2.5: Data Dive- Tonsillectomy Trends

Overview:

Tonsils (specifically the palatine tonsils) are a pair of soft tissue masses that vary in size from person to person, located adjacent to the throat (pharynx). Tonsils are part of the lymphatic organ system, which aids in fighting infections alongside the immune system. Like lymph nodes, tonsils can swell in response to infection. In the height of its popularity in the 1930's, tonsillectomies were viewed essential by the medical community because swelling was viewed as a result of disease rather than a physiological response to infection. Between the 1940's and 1970's evidence began accumulating that there was a correlation between tonsillectomies and contracting bulbar poliomyelitis (aka polio). In the 1970's, evidence then started accumulating that tonsillectomies didn't significantly decrease contracting sore throats and other throat infections as originally thought. As such, tonsillectomy rates have declined as scientific method based research guided the medical community to enforce stricter policies on when the procedure is appropriate. Figure 1.2.5.a below is indicates tonsillectomy rates in Scotland for 20+ years:

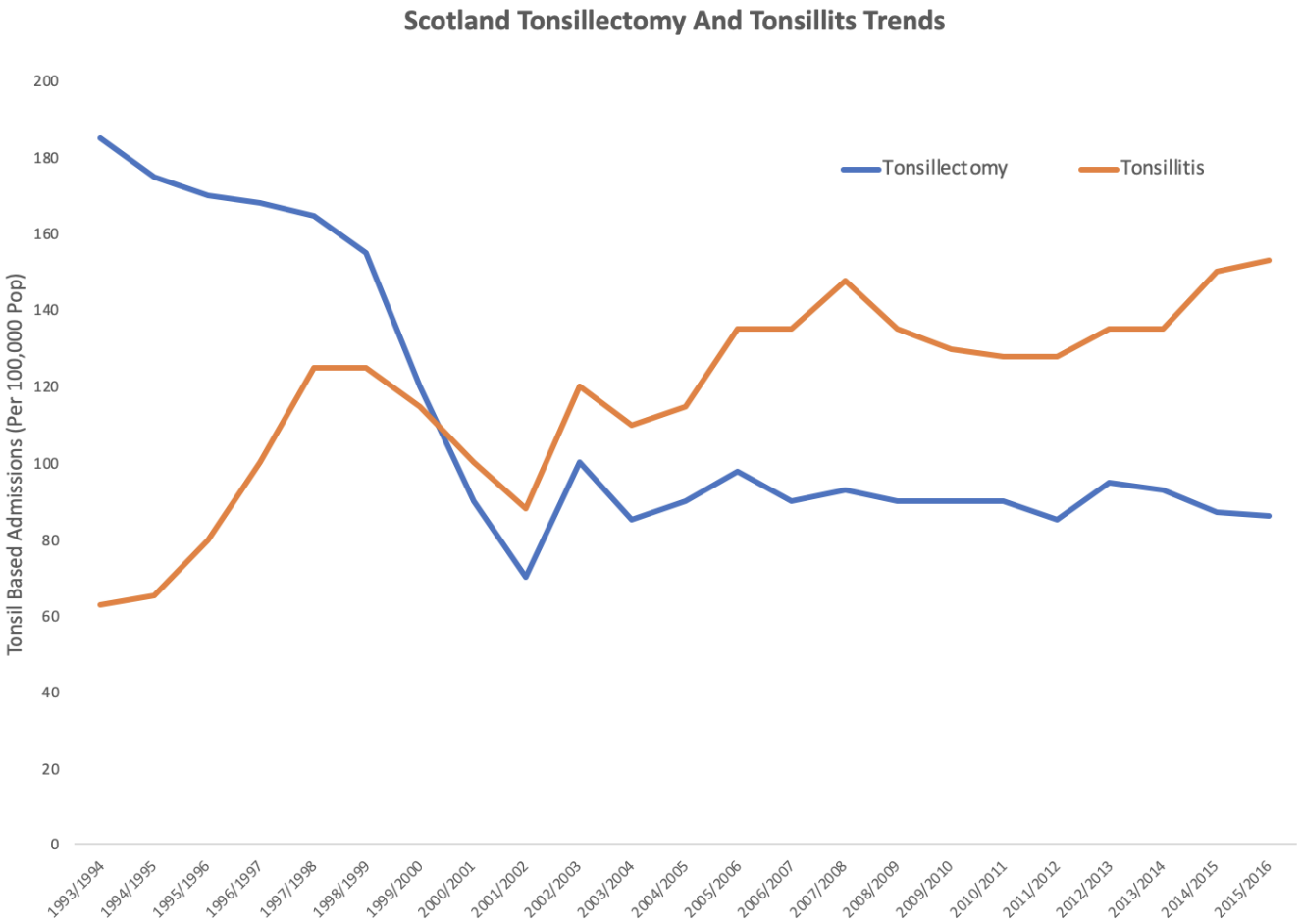


Figure 1.2.5.*a* Scotland tonsillectomy and tonsillitis trends from 1993 to 2016. Graph created by Rachel Schleiger (CC-BY-NC) modified from data in Douglas CM, Altmyer U, Cottom L, Young D, Redding P, and Clark LJ. 2019.

Questions:

1. What kind of graph is this?
2. What is the independent (explanatory) variable and the dependent (response) variable?
3. What question(s) are the authors trying to answer with this graph?
4. Describe the trend you see for tonsillectomies and tonsillitis from 1993 to 2016.

5. In 1998 new policies were put in place restricting when a tonsillectomy could take place. How did this affect the trends after 1998?
6. What does the results of this graph make you curious about? Check out this study for example.

Raw Data For Above Graph(s)

Table 1.2.5.*a* Raw data for Scotland tonsillectomy and tonsillitis trends from 1993 to 2016. Graph created by Rachel Schleiger (CC-BY-NC) modified from data in Douglas CM, Altmyer U, Cottom L, Young D, Redding P, and Clark LJ. 2019.

Years	Tonsillectomy Count	Tonsillitis Count
1993/1994	185	63
1994/1995	175	65
1995/1996	170	80
1996/1997	168	100
1997/1998	165	125
1998/1999	155	125
1999/2000	120	115
2000/2001	90	100
2001/2002	70	88
2002/2003	100	120
2003/2004	85	110
2004/2005	90	115
2005/2006	98	135
2006/2007	90	135
2007/2008	93	148
2008/2009	90	135
2009/2010	90	130
2010/2011	90	128
2011/2012	85	128
2012/2013	95	135
2013/2014	93	135
2014/2015	87	150
2015/2016	86	153

Attribution:

Rachel Schleiger (CC-BY-NC)

This page titled 1.2.5: Data Dive- Tonsillectomy Trends is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **2.5: Data Dive- Tonsillectomy Trends** by Melissa Ha and Rachel Schleiger has no license indicated.

1.2.6: Review

Summary

After completing this chapter you should be able to...

- Describe the purpose of science.
- Distinguish between objective and subjective observations.
- Distinguish between quantitative measurements and qualitative observations.
- Distinguish between inductive and deductive reasoning and relate them to descriptive and hypothesis-based science.
- Outline the steps of the scientific method and explain its cyclical nature.
- Distinguish between manipulative experiments and observational studies.
- Identify the types of variables, control group, and replicates in a scientific study.
- Discuss the importance of peer review.
- Distinguish between basic and applied science and provide examples of the value of basic science.

Science is a means of systematically gathering information about the natural world. Science is based on **objective** observations, and following the scientific method helps scientists limit **bias**. Both **induction** and **deduction** are important to the scientific method. Observations lead to a question and hypothesis, an example of inductive reasoning. Making **falsifiable predictions** based on the hypothesis and testing them through **manipulative experiments** or **observational studies** requires deductive reasoning. Finally the results are collected and scientists conclude whether the data support the hypothesis. The scientific method is a cyclical process, in which the latter steps of the process can lead back to earlier steps.

Scientists publish their findings in **scientific journals**, which require **peer review**.

Applied science focuses on solving modern problems, but **basic science** simply focuses on expanding knowledge. However, the findings of basic science can later have useful applications.

Attribution

Melissa Ha (CC-BY-NC)

This page titled 1.2.6: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **2.6: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

1.3: Math Blast- An Overview of Essential Mathematics Used in Science

Chapter Hook

Are wildfires getting more destructive and deadly than ever? It is difficult to know for sure without diving into recent and historical fire data. Currently, available data supports that wildfires are worse on both counts. As such, policy makers can utilize this knowledge to create measures to better protect lives and property. Mathematics, especially statistics, is the keystone to successful application and inquiry of the scientific method. In some cases, it also provides a high level of confidence that actions taken will not likely be a waste of time or money.



Figure 1.3.a A wildfire in a California forest. Image by Noah Berger (Public domain)

What is mathematics and statistics?

Mathematics is the abstract science of numbers, quantity and space. Mathematics may be studied in its own right (pure mathematics), or as it is applied to scientific disciplines such as (applied mathematics).

Statistics is a mathematical body of science that pertains to the collection, analysis, interpretation or explanation, and presentation of data, or as a branch of mathematics.

How does mathematics and statistics integrate into the scientific method?

Mathematics and statistics are vital for scientific inquiry. It is an especially imperative tool for disciplines that are inherently variable and require extensive collection of data. Ideally scientists would like to collect data from every individual in the population of interest, but this is rarely possible. As a result, scientists often must use data collected from a representative sample of individuals to draw inferences about basic biological phenomena for a population. This is where mathematics and statistics steps into the party.

What are the most important functions of statistics?

Some of the most important functions of statistics include:

- Description and summary of basic findings (Descriptive statistics).
- Testing hypotheses regarding relationships between variables so the cause and effect can begin to be understood (Inferential statistics).
- Presenting findings in an easily understood manner – misunderstandings keep science from moving forward!

Attribution

Rachel Schleiger (CC-BY-NC)

Math Blast

1.3.1: Getting down the basics- What are the essential mathematical and statistical terms and concepts?

1.3.2: Summarizing the data- Descriptive statistics

1.3.3: Putting it all together- Inferential statistics and hypothesis testing

1.3.4: Data Dive- California Wildfires

1.3.5: Review

This page titled 1.3: Math Blast- An Overview of Essential Mathematics Used in Science is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

1.3.1: Getting down the basics- What are the essential mathematical and statistical terms and concepts?

General terms:

- Data – Systematically recorded information.
- Value – Each measurement or observation
- Variable – The object being controlled, manipulated, measured or observed. There are two main types:
  - Independent (explanatory) – The variable that you think will affect what is being measured/observed.
  - Dependent (response) – The variable that is being measured.
- Population – Entire set of objects to be studied.
  - Parameter – Numerical characteristic of population.
- Sample – Sub-collection of objects from population.
  - Statistic – Numerical characteristic of sample from population.

Example

You are a biologist who studies how monarch butterfly (*Danaus plexippus*) populations are affected by habitat destruction. You set up a long-term study to monitor populations in degraded, intact, and restored habitats where monarchs historically have been recorded/observed.

- Population: All monarch butterflies.
- Parameter: Not possible to collect a population worth of data for monarchs. Thus, no parameters can be calculated.
- Sample: Total monarchs observed at each field site during each year of the study.
- Statistic: Any calculations/manipulations from the field site data.

Types of statistics:

- Descriptive statistics – Are calculations to summarize trends in the data. Minimally, measures of center (averages) and spread (standard deviations) from data recorded.
- Inferential statistics – The point of inferential statistics is to take data from the sample to make inferences about the population. Calculations here test hypotheses and try to find/infer cause-and-effect relationships and/or correlations.

It is important to note that statistics can only be helpful if the data from the sample is representative of the population and the interpretation of the data is unbiased!

Types of data:

- Qualitative (categorical) data – Data expressed not in terms of numbers, but rather by means of a natural language description. There are two main types of qualitative data:
  - Ordinal – When categories are in a particular order (ex: large, medium, small)
  - Nominal – When categories have no natural ordering (ex: dog breed, color)

  Graphs types used: Pie, bar
- Quantitative (numerical) data – Data expressed not by means of a natural language description, but rather in terms of numbers. There are two main types of quantitative data:
  - Continuous – Numbers where any integer or fraction can be observed (ex: time, height, or weight)
  - Discrete – A fixed number of outcomes is possible such that there are only whole integers possible (ex: counts)

  Graphs types used: Histograms, line-graphs, scatterplots

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 1.3.1: Getting down the basics- What are the essential mathematical and statistical terms and concepts? is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative)
.

- **3.1: Getting down the basics- What are the essential mathematical and statistical terms and concepts?** by Melissa Ha and Rachel Schleiger has no license indicated.

1.3.2: Summarizing the data- Descriptive statistics

How do you summarize data?

Data is summarized in two main ways: summary calculations and summary visualizations

Calculations: What types of measures are used?

To be able to interpret patterns in the data, raw data must first be manipulated and summarized into two categories of measurements: Measures of central tendency and Measures of variability. These two categories of measurements encapsulate the first step of scientific inquiry, descriptive statistics.

Measures of central tendency (center) – Provides information of how data cluster around some single middle value. There are two measures of center used most often in biological inquiry:

- **Mean** (average) – Sum of all individual values divided by total number of values in sample/population. This is the most commonly used measure of center under symmetrical distribution and is sensitive to outliers.
- **Median** – The middle value when the data set is ordered in sequential rank (highest to lowest). This is commonly used when data is skewed and is resistant to outliers.

Measures of variability (spread) – Describes how spread out or dispersed the data are. There are two main measures of spread used in biological inquiry:

- **Range** – Quantifies the distance between the largest and smallest data values.
- **Standard deviation** – Quantifies the variation or dispersion from the average of a dataset. A low standard deviation indicates that the data tends to be very close to the mean; a high standard deviation indicates that the data points are spread out over a large range of values. This calculation is sensitive to outliers.
- **Standard error** – Quantifies the variation in the means from multiple datasets or a sample distribution of your original dataset.

Visualizing the data: How are tables and graphs used?

After all desired descriptive statistics are calculated, they are typically visually summarized into either a table or graph.

Tables:

A **table** is a set of data values arranged into columns and rows. Typically the columns encompass a broad data category, and the rows encompass another. Within each broad category there are subcategories that determine how many columns and rows the table consists of. Tables are used to both collect and summarize data. However, most of the time when tables are presented, they consist summarized data, not raw data. Although tables allow summarized data to be presented in an orderly manner, most people prefer to translate tables into the more powerful data visualization tool, a graph.

Graphs:

A **graph** is a diagram showing the relation between variable quantities, typically of two variables, each measured along one of a pair of axes at right angles. Graphs can look like a chart or drawing. Most graphs use bars, lines, or parts of a circle to display data. However, there are sometimes when graphs are overlaid on top of maps to also display geographical location, or are even animated to be interactive.

Major graph type categories:

- **Circle/Pie** – A circular chart divided into slices to illustrate numerical proportion. In a pie chart, the arc length of each slice (and consequently its central angle and area), is proportional to the quantity it represents. While it is named for its resemblance to a pie which has been sliced, there are variations on the way it can be presented.
- **Line** – A type of chart which displays information as a series of data points called 'markers' connected by straight line segments. It is a basic type of chart common in many fields. It is similar to a scatter plot except that the measurement points are ordered (typically by their x-axis value) and joined with straight line segments. A line chart is often used to visualize a trend in data over intervals of time – a time series – thus the line is often drawn chronologically.

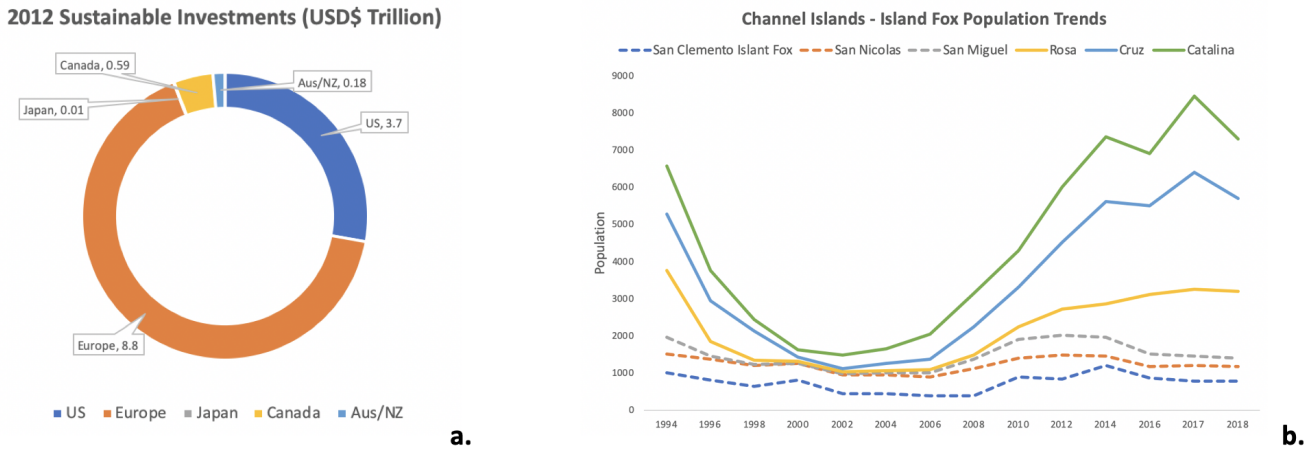


Figure 1.3.2.*a*: Examples of a circle/pie graph (a.) and a line graph (b.). Image created by Rachel Schleiger (CC-BY-NC).

- **Scatter plot** – Is a graph in which the values of two variables are plotted along the horizontal and vertical axes, the pattern of the resulting points revealing any correlation preset. The data are displayed as a collection of points, each having the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis.

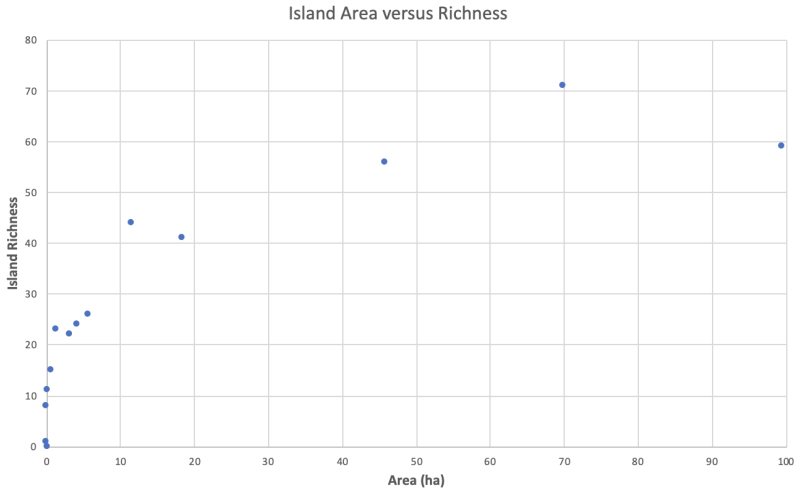


Figure 1.3.2.*b*: Example of a scatter plot. Image created by Rachel Schleiger (CC-BY-NC).

- **Bar** – A chart or graph that presents categorical data with rectangular bars with heights or lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally.
- **Histogram** – Is an approximate representation of the distribution of numerical data. To construct a histogram, the first step is to "bin" (or "bucket") the range of values—that is, divide the entire range of values into a series of intervals—and then count how many values fall into each interval. The bins are usually specified as consecutive, non-overlapping intervals of a variable. The bins (intervals) must be adjacent (meaning there are not spaces between them like there are in bar graphs), and are often (but not required to be) of equal size. If the bins are of equal size, a rectangle is erected over the bin with height proportional to the frequency—the number of cases in each bin.

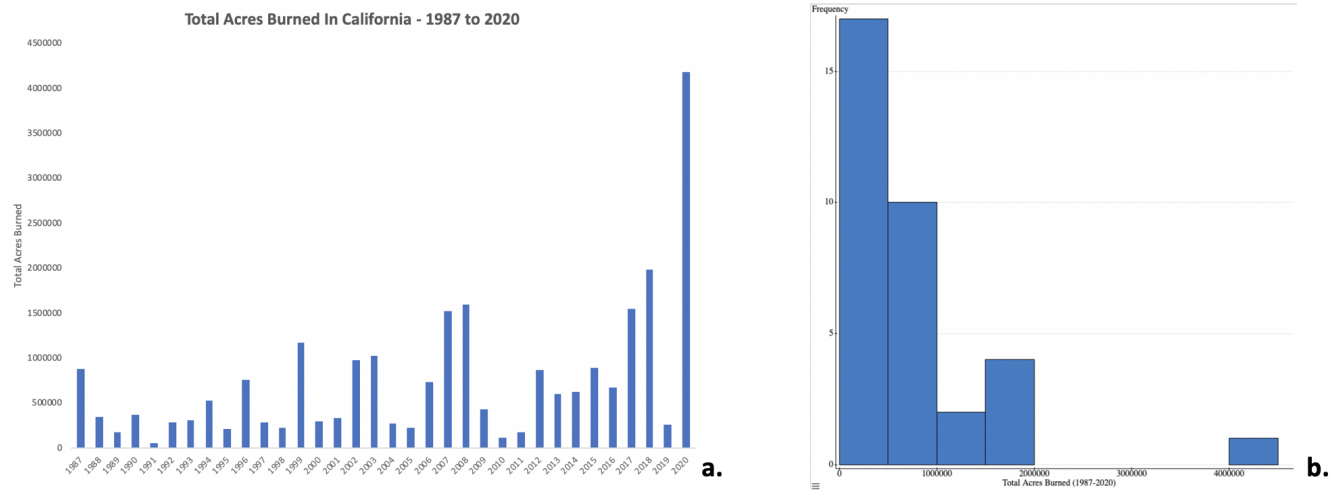


Figure 1.3.2.*c*: Examples of a bar graph (a.) and a histogram (b.) using the same dataset. Image created by Rachel Schleiger (CC-BY-NC).

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 1.3.2: Summarizing the data- Descriptive statistics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **3.2: Summarizing the data- Descriptive statistics** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

1.3.3: Putting it all together- Inferential statistics and hypothesis testing

What is a hypothesis and are there different kinds?

Biological (Scientific) hypothesis: An idea that proposes a tentative explanation about a phenomenon or a narrow set of phenomena observed in the natural world. This is the backbone of all scientific inquiry! As such it is important to have a solid biological hypothesis before moving forward in the scientific method (i.e. procedures, results, discussion). After the creation of a solid biological hypothesis, it can then be simplified into a statistical hypothesis (as defined below) that will become the basis for how the data will be analyzed and interpreted.

Statistical hypotheses: After defining a strong biological hypothesis, a statistical hypothesis can be created based on what you will predict will be the measured outcome(s) (dependent variable(s)). If a study has multiple measured outcomes there can be multiple statistical hypotheses. Each statistical hypothesis will have two components (Null and Alternative).

- **Null hypothesis** (Ho) –This hypothesis states that there is no relationship (or no pattern) between the independent and dependent variables.
- **Alternative hypothesis** (H1) – This hypothesis states that there is a relationship (or is a pattern) between the independent and dependent variables.

Independent versus dependent variables: For both biological and statistical hypotheses there should be two basic variables defined:

- **Independent (explanatory) variable** – It is usually what phenomena you think will affect the measure you are interested in (dependent variable).
- **Dependent (response) variable** – A dependent variable is what you measure in the experiment and what is affected during the experiment. The dependent variable responds to (depends on) the independent variable. In a scientific experiment, you cannot have a dependent variable without an independent variable.

Example

Yellow-billed Cuckoo nests were counted during breeding season in degraded, restored, and intact riparian habitats to see overall habitat preference for nesting sites increased with habitat health.

- Scientific hypothesis: Yellow-billed Cuckoo will have habitat preferences because of habitat health/status.
- Statistical hypotheses: (Ho) There will be no differences in number of nests between habitats with different health/status. (H1) There will be more nests in restored and intact habitats compared to degraded.
- Independent variable = Habitat health/status
- Dependent variable = Number of nests counted

How do you make conclusions?

Finally, after defining the biological hypothesis, statistical hypothesis, and collecting all your data, a researcher can begin statistical analysis. A statistical test will mathematically “test” your data against the statistical hypothesis. The type of statistical test that is used depends on the type and quantity of variables in the study, as well as the question the researcher wants to ask. After computing the statistical test, the outcome will indicate which statistical hypothesis is more likely. This, in turn indicates to scientists what level of inference can be gained from the data compared to the biological hypothesis (the focus point of the study). Then a conclusion can be made based on the sample about the entire population. It is important to note that the process does not stop here. Scientists will want to continue to test this conclusion until a clear pattern emerges (or not) or to investigate similar but different questions.

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 1.3.3: Putting it all together- Inferential statistics and hypothesis testing is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **3.3: Putting it all together- Inferential statistics and hypothesis testing** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

1.3.4: Data Dive- California Wildfires

Overview

When the United States Forest Service was established in 1905 it was declared that, “Today we understand that forest fires are wholly within the control of men.” Humanity assumes time and time again that nature can be controlled, and the National Forest Service were shown how wrong their assumption was 5 years later… The Great Fire of 1910 burned about 3 million acres and killed 86 people. As a consequence of this fire, serious fire suppression policies were put in place and Smokey The Bear was slapped onto posters to represent the cause. At that point, Smokey the Bear told us, "only you can prevent forest fire," indicating that all fire is bad. It has only been in the last couple decades have these policies been questioned and the repercussions of fire suppression of them realized. Now Smokey has updated the saying to, "only you can prevent wildfire," indicating that only some types of fire are bad. The figure below illustrates the number of acres burned in California between 1987 and 2020:

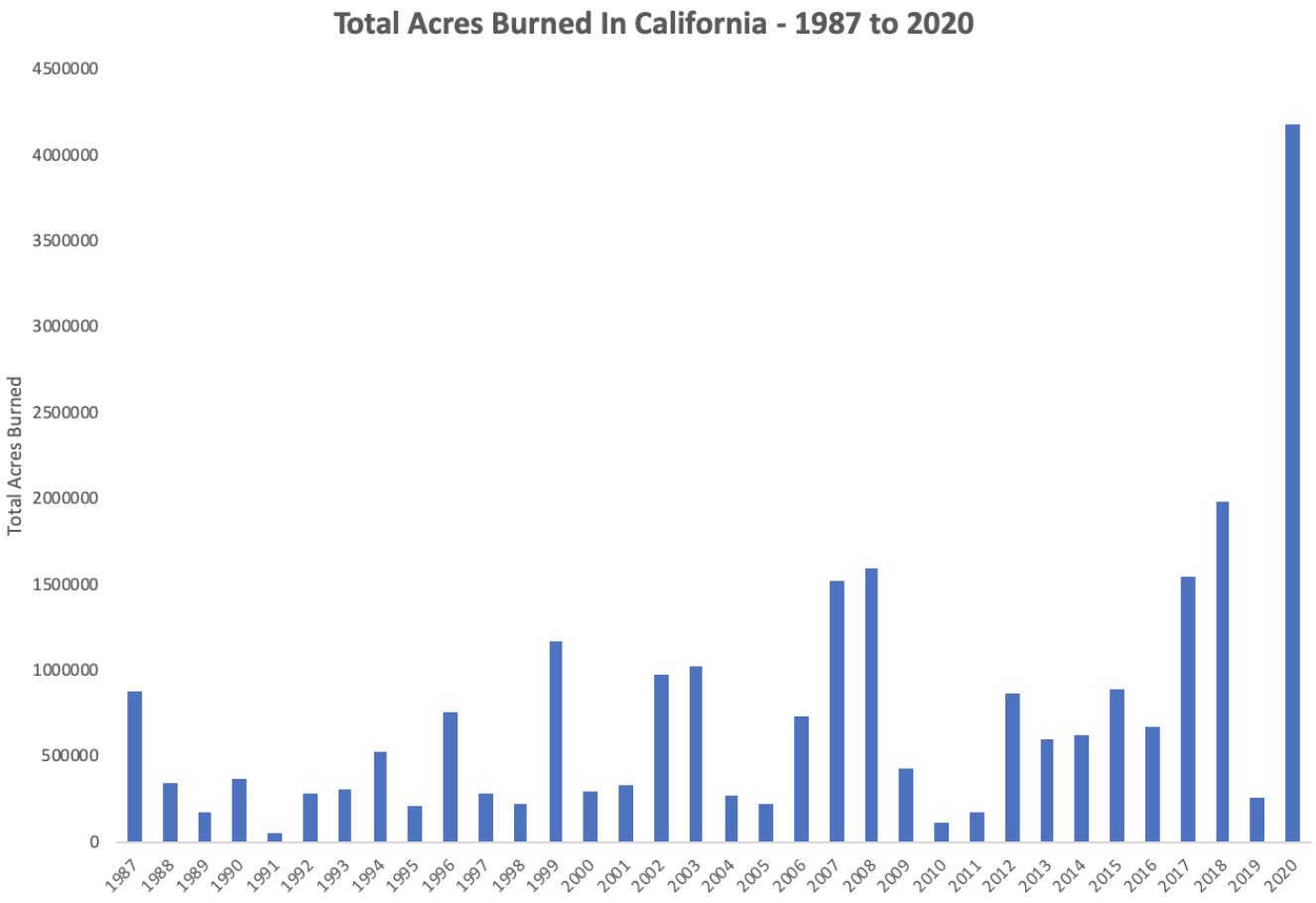


Figure 1.3.4.*a*: Total acres burned through the years in California. Graph created by Rachel Schleiger (CC-BY-NC) modified from data collected by CalFire

Questions

1. What kind of graph is this?
2. What is the independent (explanatory) variable and the dependent (response) variable?
3. What question(s) are the authors trying to answer with this graph?
4. What trend(s) can be observed in this graph? Support your answer by referring to appropriate patterns in the graph.

5. How can we use the results of this graph to inform future fire policies/projects/etc?
6. What does the results of this graph make you curious about?

Raw Data For Above Graph(s)

Table 1.3.4.*a***:** Raw data table of total acres burned through the years in California. Table by Rachel Schleiger (CC-BY-NC) modified from data collected by CalFire

Year	Total Acres Burned
1987	873000
1988	345000
1989	173400
1990	365200
1991	44200
1992	282745
1993	309779
1994	526219
1995	209815
1996	752372
1997	283885
1998	215412
1999	1172850
2000	295026
2001	329126
2002	969890
2003	1020460
2004	264988
2005	222538
2006	736022
2007	1520362
2008	1593690
2009	422147
2010	109529
2011	168545
2012	869599
2013	601635
2014	625540
2015	893362
2016	669534
2017	1548429
2018	1975086
2019	259823

Year	Total Acres Burned
2020	4177856

Attributions

Rachel Schleiger (CC-BY-NC)

This page titled 1.3.4: Data Dive- California Wildfires is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **3.4: Data Dive- California Wildfires** by Melissa Ha and Rachel Schleiger has no license indicated.

1.3.5: Review

Summary

After completing this chapter you should be able to...

- Know how mathematics and statistics fits into the scientific method
- Define the basic mathematical terms
- Know what data is what what types of data there are
- Understand the difference between descriptive and inferential statistics
- Have a general idea of what types of graphs there are to help visualize data
- Understand the difference between scientific and statistical hypotheses
- Know the process for how conclusions are made

Statistics are based on samples of data collected to be representative of a population of interest due to the impossibility of collecting a complete dataset. There are two types of statistics, **descriptive** (summary calculations) and **inferential** (hypothesis testing). **Data** can come from anywhere! There are two main types of data, **qualitative** (categorical) and **quantitative** (numerical).

For descriptive statistics, measures of **center** and **spread** are commonly calculated in addition to some sort of visualization of the data. The type of calculations and visualizations used depend on the study parameters and objectives.

For inferential statistics, **hypotheses** are clearly defined and analyses are run. There are two main types of hypotheses, **scientific** and **statistical**. The scientific proposes a potential explanation about a phenomenon while a statistical is used to predictions about measured outcomes and has two forms, **null** and **alternative**. **Independent** and **dependent** variables are used to structure the ideas outlined in the hypotheses. Once a study is performed to test the hypotheses then conclusions can be drawn based on evidence identified for each hypothesis.

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 1.3.5: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **3.5: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

CHAPTER OVERVIEW

2: Environmental Chemistry

Environmental chemistry is an interdisciplinary science that includes atmospheric, aquatic and soil chemistry, as well as heavily relying on analytical chemistry and being related to environmental and other areas of science. Environmental chemistry is the study of chemical processes occurring in the environment which are impacted by humankind's activities.

- 00: Front Matter
  - TitlePage
  - InfoPage
  - Table of Contents
  - Licensing
- Geochemistry
  - Licensing
  - 2.1.1: The Earth and its Lithosphere
    - 2.1.1.1: Composition and Structure of the Earth
    - 2.1.1.2: Origin of the Elements
    - 2.1.1.3: Formation and evolution of the Earth
    - 2.1.1.4: The Earth's crust
  - 2.1.2: The Hydrosphere
    - 2.1.2.1: Water, Water Everywhere...
    - 2.1.2.2: The hydrosphere and the oceans
    - 2.1.2.3: Chemistry and geochemistry of the oceans
    - 2.1.2.4: Chemical budgets of oceanic elements
  - 2.1.3: The Atmosphere
    - 2.1.3.1: Structure and Composition of the Atmosphere
    - 2.1.3.2: Origin and Evolution of the Atmosphere
  - 2.1.4: The Biosphere
    - 2.1.4.1: Chemistry and Energetics of the Life Process
    - 2.1.4.2: Biogeochemical Evolution
    - 2.1.4.3: Gaia - Bioregulation of the Environment
- 2.2: Supplemental Modules (Environmental Chemistry)
  - 2.2.1: Acid Rain
    - 2.2.1.1: Acid Rain
    - 2.2.1.2: Acid Rain Transport
    - 2.2.1.3: Acid Snow
    - 2.2.1.4: Electricity Generation
    - 2.2.1.5: Playground 1
    - 2.2.1.6: Sources of Nitrogen Oxides
    - 2.2.1.7: Sources of Sulfur Oxides
  - 2.2.2: Aquatic Chemistry
    - 2.2.2.1: 1. Acid-Base Chemistry of Natural Aquatic Systems
    - 2.2.2.2: 2. Carbonate Equilibria in Natural Waters

2.2.2.3: 3. Redox Equilibria in Natural Waters
2.2.2.4: 4. Solids in Contact With Natural Waters
2.2.2.5: Fundamental Characteristics of Water
2.2.2.6: Natural Water
2.2.2.7: Water Biology
2.2.2.8: Water Chemistry
2.2.2.9: Water Physics
2.2.2.10: Water Treatment
2.2.3: Atmospheric Chemistry
2.2.3.1: Atmosphere
2.2.3.2: Carbon Cycle
2.2.3.3: Carbon Oxides
2.2.3.4: Ozone
2.2.3.5: Photosynthesis
2.2.4: Biochemical Cycles
2.2.4.1: Biochemical Cycles
2.2.4.2: Carbon Cycle
2.2.4.3: Nitrogen Cycle
2.3: Environmental Toxicology (van Gestel et al.)
00: Front Matter
Table of Contents
2.3.1: Environmental Toxicology
2.3.1.1: Environmental toxicology
2.3.1.2: DPSIR
2.3.1.3: Short history
2.3.2: Environmental Chemistry, Chemicals
2.3.2.1: Introduction
2.3.2.2: Pollutants with specific properties
2.3.2.3: Pollutants with specific use
2.3.3: Environmental Chemistry - From Fate to Exposure
2.3.3.1: Environmental compartments
2.3.3.2: Sources of chemicals
2.3.3.3: Pathways and processes determining chemical fate
2.3.3.4: Partitioning and Partitioning Constants
2.3.3.5: Metal Speciation
2.3.4: Toxicology
2.3.4.1: Toxicokinetics
2.3.4.2: Toxicodynamics and Molecular Interactions
2.3.4.3: Toxicity Testing
2.3.4.4: Increasing ecological realism in toxicity testing
2.3.4.5: Availability and bioavailability
2.3.4.6: Degradation
2.3.4.7: Modelling exposure
2.3.4.8: Section 4.1. Toxicokinetics

2.3.4.9: Section 4.4. Increasing ecological realism in toxicity testing
2.3.4.10: Section 4.2. Toxicodynamics and Molecular Interactions
2.3.4.11: Section 4.3. Toxicity testing
2.3.5: Population, Community and Ecosystem Ecotoxicology
2.3.5.1: Introduction- Linking population, community and ecosystem responses
2.3.5.2: Population ecotoxicology in laboratory settings
2.3.5.3: Wildlife population ecotoxicology
2.3.5.4: Trait-based approaches
2.3.5.5: Population models
2.3.5.6: Metapopulations
2.3.5.7: Community ecotoxicology
2.3.5.8: Structure versus function incl. ecosystem services
2.3.5.9: Landscape ecotoxicology
2.3.6: Risk Assessment and Regulation
2.3.6.1: Introduction - The Essence of Risk Assessment
2.3.6.2: Ecosystem services and protection goals
2.3.6.3: Predictive risk assessment approaches and tools
2.3.6.4: Diagnostic risk assessment approaches and tools
2.3.6.5: Regulatory Frameworks
2.3.6.6: Risk management and risk communication
2.3.6.7: Risk perception
Key Elements of Green Chemistry
Licensing
2.4.1: Key Elements of Green Chemistry
2.4.1.1: Introduction to Green Chemistry
2.4.1.2: Green Chemistry Concepts
2.4.1.3: Case Study
2.4.1.4: Conclusions and Review
2.4.2: Life-Cycle Analysis
2.4.2.1: Introduction to Life Cycles
2.4.2.2: LCA/LCIA Concepts
2.4.2.3: Conclusions and Review
2.4.3: Hazards
2.4.3.1: Introduction to Hazards of Chemistry
2.4.3.2: Hazard Concepts
2.4.3.3: Case Study - Badger Army Ammunition Plant
2.4.3.4: Green Technologies for Safer Chemical Production
2.4.3.5: Conclusions and Review Questions
2.4.4: Alternative Solvents
2.4.4.1: Introduction to Solvents
2.4.4.2: Solvent Concepts
2.4.4.3: Solvent Alternatives
2.4.4.4: Conclusions and Review
2.4.5: Alternative Reagants

2.4.5.1: Introduction to Reagants
2.4.5.2: Reagant Concepts
2.4.5.3: Specific Replacements
2.4.5.4: Conclusions and Review
2.4.6: Reaction Types, Design, and Efficiency
2.4.6.1: Introduction to Reactions
2.4.6.2: Reaction Concepts
2.4.6.3: Reaction Design Concepts
2.4.6.4: Conclusions and Review
Detailed Licensing
Green Chemistry and the Ten Commandments of Sustainability
Licensing
2.5.1: Sustainability and the Environment
2.5.1.1: Sustainability
2.5.1.2: The Environment and the Five Environmental Sphere
2.5.1.3: Seeing Green
2.5.1.4: Natural Capital of the Earth
2.5.1.5: Sustainability as a Group Effort- It Takes a (Very Big) Village
2.5.1.6: Sustainable Energy- Away from the Sun and Back Again
2.5.1.7: Green Science
2.5.1.8: Green Technology
2.5.1.9: Sustainability and the Eco-Economy
2.5.1.10: Green Products and Services- Design for Sustainability
2.5.1.11: Questions and Problems
2.5.2: The Key Role of Chemistry and Making Chemistry Green
2.5.2.1: Chemistry is Good (and Unavoidable)
2.5.2.2: The Environment and its Chemistry
2.5.2.3: What is Environmental Chemistry?
2.5.2.4: Environmental Pollution
2.5.2.5: Practice of Green Chemistry
2.5.2.6: Green Chemistry and Synthetic Chemistry
2.5.2.7: Reduction of Risk- Hazard and Exposure
2.5.2.8: The Risks of No Risks
2.5.2.9: Waste Prevention
2.5.2.10: Basic Principles of Green Chemistry
2.5.2.11: Some Things to Know About Chemistry before You Even Start
2.5.2.12: Combining Atoms to Make Molecules and Compounds
2.5.2.13: The process of making and breaking chemical bonds- chemical reactions
2.5.2.14: The Nature of Matter and States of Matter
2.5.2.15: References and Questions
2.5.3: The Elements - Basic Building Blocks of Green Chemicals
2.5.3.1: Elements, Atoms, and Atomic Theory
2.5.3.2: Hydrogen - The Simplest Atom
2.5.3.3: Helium - The First Noble Gas

2.5.3.4: Lithium, The First Metal
2.5.3.5: The Second Period of the Periodic Table
2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons
2.5.3.7: Completing the 20-Element Periodic Table
2.5.3.8: The Brief Periodic Table is Complete
2.5.3.E: The Elements - Basic Building Blocks of Green Chemicals (Exercises)
2.5.3.R: References
2.5.4: Compounds- Safer Materials for a Safer World
2.5.4.1: Chemical Bonds and Compound Formation
2.5.4.2: Electrons Involved in Chemical Bonds and Octets of Electrons
2.5.4.3: Sodium Chloride and Ionic Bonds
2.5.4.4: Covalent Bonds in H2 and Other Molecules
2.5.4.5: Covalent Bonds in Compounds
2.5.4.6: Covalent Bonds and Green Chemistry
2.5.4.7: Predicting Formulas of Covalently Bound Compounds
2.5.4.8: Chemical Formulas, the Mole, and Percentage Composition
2.5.4.9: What Are Chemical Compounds Called?
2.5.4.10: Acids, Bases and Salts
2.5.4.11: Questions and Problems
2.5.5: Chemical Reactions- Making Materials Safely and Sustainable
2.5.5.1: Describing What Happens with Chemical Equations
2.5.5.2: Balancing Chemical Equations
2.5.5.3: Just Because You Can Write It Doesn't Mean That It Will Happen
2.5.5.4: Yield And Atom Economy in Chemical Reactions
2.5.5.5: Catalysts That Make Reactions Go
2.5.5.6: Kinds of Chemical Reactions
2.5.5.7: Oxidation-Reduction Reactions and Green Chemistry
2.5.5.8: Quantitative Information from Chemical Reactions
2.5.5.9: Energy in Chemical Reactions
2.5.5.10: Stoichiometry by the Mole Ratio Method
2.5.5.11: Limiting Reactant and Percent Yield
2.5.5.12: Titrations - Measuring Moles by Volume of Solution
2.5.5.13: Industrial Chemical Reactions - The Solvay Process
2.5.5.14: Questions and Problems
2.5.6: The Wonderful World of Carbon - Organic Chemistry and Biochemicals
2.5.6.1: Rings and Chains of Carbon Atoms
2.5.6.2: Compounds of Carbon and Hydrogen - Hydrocarbons
2.5.6.3: Using Lines To Show Organic Structural Formulas
2.5.6.4: Functional Groups
2.5.6.5: Giant Molecules from Small Organic Molecules
2.5.6.6: Questions and Problems
2.5.6.7: Supplementary References
2.5.7: Chemistry of Life and Green Chemistry
2.5.7.1: Green Biochemistry

2.5.7.2: Biochemistry and the Cells
2.5.7.3: Carbohydrates
2.5.7.4: Proteins
2.5.7.5: Lipids - Fats, Oils, and Hormones
2.5.7.6: Nucleic Acids
2.5.7.7: Enzymes
2.5.7.8: Biochemical Processes in Metabolism
2.5.7.9: Biochemistry of Toxic Substances and Toxicological Chemistry
2.5.7.10: Literature Cited
2.5.7.11: Questions and Problems
2.5.7.12: Supplementary References
2.5.8: The Five Environmental Spheres and Biogeochemical Cycle
2.5.8.1: Five Environmental Spheres
2.5.8.2: The Hydrosphere
2.5.8.3: The Atmosphere
2.5.8.4: The Geosphere
2.5.8.5: The Biosphere
2.5.8.6: The Anthrosphere
2.5.8.7: Cycles of Matter
2.5.8.8: Questions and Problems
2.5.8.9: Supplementary References
2.5.9: Water - the Ultimate Green Substance
2.5.9.1: H_2O - Simple Formula, Remarkable Molecule
2.5.9.2: Occurrence, Availability , and Utilization of Water on Earth
2.5.9.3: Water Chemistry
2.5.9.4: Water Pollution
2.5.9.5: Greening of Water - Purification Before Use
2.5.9.6: Wastewater Treatment
2.5.9.7: Advanced Wastewater Treatment and Recycling
2.5.9.8: The Many Uses of Water
2.5.9.9: Hot Water - Pressurized Subcritical Water
2.5.9.10: Supercritical Water
2.5.9.11: Literature Cited
2.5.9.12: Questions and Problems
2.5.9.13: Supplementary References
2.5.10: Blue Skies for a Green Environment
2.5.10.1: More Than Just Air to Breathe
2.5.10.2: The Gas Laws
2.5.10.3: The Protective Atmosphere
2.5.10.4: Atmospheric Chemistry and Photochemical Reactions
2.5.10.5: The Enormous Importance of Climate
2.5.10.6: Atmospheric Particle Pollutants
2.5.10.7: Pollutant Gaseous Oxides
2.5.10.8: Acid Rain

2.5.10.9: Miscellaneous Gases in the Atmosphere
2.5.10.10: Photochemical Smog
2.5.10.11: Natural Capital of the Atmosphere
2.5.10.12: Energy and Mass Transfer in the Atmosphere
2.5.10.13: Questions and Problems
2.5.10.14: Supplementary References
2.5.11: The Geosphere and a Green Earth
2.5.11.1: Probing Deep into the Geosphere - A Well Too Far?
2.5.11.2: The Nature of the Geosphere
2.5.11.3: The Geosphere as a Source of Natural Capital
2.5.11.4: Environmental Hazards of the Geosphere
2.5.11.5: Water on and in the Geosphere
2.5.11.6: Anthrospheric Influences on the Geosphere
2.5.11.7: The Geosphere as a Waste Repository
2.5.11.8: Have You Thanked a Clod Today?
2.5.11.9: Plant Nutrients and Fertilizers in Soil
2.5.11.10: Soil and Plants Related To Wastes and Pollutants
2.5.11.11: Soil Loss - Desertification and Deforestation
2.5.11.12: Literature Cited and Supplementary References
2.5.11.13: Questions and Problems
2.5.12: The Biosphere and the Role of Green Chemistry in Feeding a Hungry World
2.5.12.1: Pigweed’s Revenge
2.5.12.2: The Biosphere
2.5.12.3: Cells - Basic Units of Life
2.5.12.4: Metabolism and Control in Organisms
2.5.12.5: Reproduction and Inherited Traits
2.5.12.6: Stability and Equilibrium of the Biosphere
2.5.12.7: DNA and the Human Genome
2.5.12.8: Genetic Engineering
2.5.12.9: Biological Interaction with Environmental Chemicals
2.5.12.10: Biodegradation
2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture
2.5.12.12: Agricultural Applications of Genetically Modified Organisms
2.5.12.13: The Anthrosphere in Support of the Biosphere
2.5.12.14: Livestock and their Wastes
2.5.12.15: Questions and Problems
2.5.12.16: Supplementary References
2.5.13: The Anthrosphere, Industrial Ecology, and Green Chemistry
2.5.13.1: Industrial Ecology and Industrial Ecosystems
2.5.13.2: Metabolic Processes in Industrial Ecosystems
2.5.13.3: Life Cycles in Industrial Ecosystems
2.5.13.4: Kinds of Products
2.5.13.5: Attributes Required by an Industrial Ecosystem
2.5.13.6: The Kalundborg Industrial Ecosystem

2.5.13.7: Environmental Impacts of Industrial Ecosystems
2.5.13.8: Green Chemistry and Industrial Ecology
2.5.13.9: Predicting and Reducing Hazards
2.5.13.10: The E-Factor in Green Chemistry
2.5.13.11: Catalysts and Catalysis
2.5.13.12: Biocatalysis with Enzymes
2.5.13.13: Energizing Chemical Reactions and Process Intensification
2.5.13.14: Solvents and Alternate Reaction Media
2.5.13.15: Feedstocks and Reagents
2.5.13.16: Literature Cited and Supplementary References
2.5.13.17: Questions and Problems
2.5.14: Feeding the Anthrosphere- Utilizing Renewable and Biological Materials
2.5.14.1: Feeding the Anthrosphere
2.5.14.2: Utilization of Feedstocks
2.5.14.3: Biological Feedstocks
2.5.14.4: Biological Sources of Chemicals
2.5.14.5: Biorefineries and Biomass Utilization
2.5.14.6: Monosaccharide Feedstocks - Glucose, Fructose, and Xylose
2.5.14.7: Cellulose
2.5.14.8: Feedstocks from Cellulose Wastes
2.5.14.9: Lignin
2.5.14.10: Direct Biosynthesis of Polymers
2.5.14.11: Bioconversion Processes for Synthetic Chemicals
2.5.14.12: Bamboo - Ancient Material for the Future
2.5.14.13: Literature Cited
2.5.14.14: Questions and Problems
2.5.15: Sustainable Energy- The Essential Basis of Green Systems
2.5.15.1: Sustainability and Energy
2.5.15.2: What is Energy?
2.5.15.3: Radiant Energy from the Sun
2.5.15.4: Sources of Energy Used in the Anthrosphere
2.5.15.5: Conversions Between Forms of Energy
2.5.15.6: Green Technology for Energy Conversion
2.5.15.7: Energy Conservation
2.5.15.8: Depletable Fossil Fuels
2.5.15.9: Carbon Sequestration for Fossil Fuel Utilization
2.5.15.10: Nuclear Energy
2.5.15.11: Renewable Energy Sources - Solar Energy
2.5.15.12: Energy from Wind and Water
2.5.15.13: Biomass Energy
2.5.15.14: Geothermal Energy
2.5.15.15: Hydrogen for Energy Storage and Utilization
2.5.15.16: Literature Cited and Supplementary References
2.5.15.17: Questions and Problems

2.5.16: Terrorism, Toxicity, and Vulnerability- Green Chemistry and Technology in Defense of Human Welfare
2.5.16.1: Vulnerability to Terrorist Attack
2.5.16.2: Protecting the Anthrosphere
2.5.16.3: Substances That Explode, Burn, or React Violently
2.5.16.4: Toxic Substances and Toxicology
2.5.16.5: Toxic Chemical Attack
2.5.16.6: Protecting Water, Food, and Air
2.5.16.7: Detecting Hazards
2.5.16.8: Green Chemistry to Combat Terrorism
2.5.16.9: Green Chemistry for Sustainable Prosperity and a Safer World
2.5.16.10: Green Chemistry to Combat Terrorism
2.5.16.11: Green Chemistry for Sustainable Prosperity and and a Safer World
2.5.16.12: Literature Cited
2.5.16.13: Questions and Problems
2.5.17: The Ten Commandments of Sustainability and Sensible Measure
2.5.17.1: We Cannot Go on Like This
2.5.17.2: The First Commandment
2.5.17.3: The Second Commandment
2.5.17.4: The Third Commandment
2.5.17.5: The Fourth Commandment
2.5.17.6: The Fifth Commandment
2.5.17.7: The Sixth Commandment
2.5.17.8: The Seventh Commandment
2.5.17.9: The Ninth Commandment
2.5.17.10: The Tenth Commandment
2.5.17.11: Some Sensible Measures for Sustainability
2.5.17.12: The Eighth Commandment
2.5.17.13: Literature Cited
2.5.17.14: Questions and Problems
2.5.17.15: Supplementary References
Detailed Licensing
Toxicology
Licensing
Preface
2.6.1: Pathophysiology
2.6.1.1: What is Pathophysiology?
2.6.1.2: Outcomes of Targeted and Non-Targeted Toxicity
2.6.1.3: Cellular Response to Toxicant-Induced Injury
2.6.1.4: Repair and Adaptation
2.6.1.5: Patterns of Toxic Injury
2.6.1.6: Final Evaluation
2.6.2: Biochemistry and Molecular Genetics
2.6.2.1: Introduction to Biomolecules and Cell Components
2.6.2.2: Cell Structure and Subcellular Compartments

2.6.2.3: DNA and RNA Metabolism
2.6.2.4: Epigenetic Mechanisms
2.6.2.5: Section 2 Final Evaluation
2.6.3: Principles of Genetic Toxicology
2.6.3.1: Introduction to Genetic-toxicology Assay
2.6.3.2: Different Genetic Damages or Mutations
2.6.3.3: Different Genetic-Toxicology Assays
2.6.3.4: Different Cytotoxicity Assays
2.6.3.5: Epigenetics Assay
2.6.3.6: Section 3 Final Evaluation
2.6.4: Applied Systems Toxicology
2.6.4.1: Systems Toxicology
2.6.4.2: Dose Level and Applied Toxicology
2.6.4.3: Tools and Technologies in Systems Toxicology
2.6.4.4: Other Approaches for Predictive Toxicity Modeling
2.6.4.5: Technologies Used In Systems Biology/Toxicology
2.6.4.6: Takeaways Summary
2.6.4.7: Section 4 Final Evaluation
2.6.5: Regulatory Toxicology
2.6.5.1: Introduction to Regulatory Toxicology
2.6.5.2: Global Regulatory Toxicology
2.6.5.3: Topic 3- Regional Regulatory Toxicology
2.6.5.4: National Regulatory Toxicology
2.6.5.5: State Regulatory Toxicology
2.6.5.6: Non-Governmental Regulatory Toxicology
2.6.5.7: Section 5 Final Evaluation
2.6.6: Principles of Toxicology
2.6.6.1: Introduction to Toxicology
2.6.6.1.1: What is toxicology?
2.6.6.1.2: Basic Terminology
2.6.6.2: Dose and Dose Response
2.6.6.2.1: Dose and It's Impact on Toxicity
2.6.6.2.2: The Dose Response Relationship
2.6.6.2.3: Dose Estimates of Toxic Effects
2.6.6.2.4: Determining the Safety of a Drug
2.6.6.2.5: NOAEL and LOAEL
2.6.6.3: Toxic Effects
2.6.6.3.1: Types of Toxic Effects
2.6.6.3.2: Factors Affecting Toxicity
2.6.6.3.3: Systemic Toxic Effects
2.6.6.3.4: Organ Specific Toxic Effects
2.6.6.4: Interactions
2.6.6.4.1: Interactions
2.6.6.5: Toxicity Testing Methods

2.6.6.5.1: Testing and Assessing Toxicity
2.6.6.5.2: Clinical Investigations and Other Types of Human Data
2.6.6.5.3: Epidemiology Studies
2.6.6.6: Risk Assessment
2.6.6.6.1: Risk Assessment
2.6.6.6.2: Hazard Identification
2.6.6.6.3: Dose-Response Assessment
2.6.6.6.4: Exposure Assessment
2.6.6.6.5: Risk Characterization
2.6.6.7: Exposure Standards and Guidelines
2.6.6.7.1: Exposure Standards and Guidelines
2.6.6.7.2: Regulation of Consumer Products and Drug Safety
2.6.6.7.3: Environmental Exposure Standards/Guidelines
2.6.6.7.4: Occupational (Workplace) Exposure Standards/Guidelines/Approaches
2.6.6.8: Basic Physiology
2.6.6.8.1: Introduction to Basic Physiology
2.6.6.8.2: Homeostasis
2.6.6.8.3: Organs and Organ Systems
2.6.6.8.4: Tissues
2.6.6.8.5: Cells
2.6.6.8.6: Chemicals
2.6.6.9: Introduction to Toxicokinetics
2.6.6.9.1: What is Toxicokinetics
2.6.6.10: Absorption
2.6.6.10.1: Introduction to Absorption
2.6.6.10.2: Gastrointestinal Tract
2.6.6.10.3: Respiratory Tract
2.6.6.10.4: Dermal Route
2.6.6.10.5: Other Routes of Exposure
2.6.6.11: Distribution
2.6.6.11.1: Introduction to Distribution
2.6.6.11.2: Influence of Route of Exposure
2.6.6.11.3: Disposition Models
2.6.6.11.4: Structural Barriers to Distribution
2.6.6.11.5: Storage Sites
2.6.6.12: Biotransformation
2.6.6.12.1: Introduction to Biotransformation
2.6.6.12.2: Chemical Reactions
2.6.6.12.3: Biotransformation Sites
2.6.6.12.4: Modifiers of Biotransformation
2.6.6.13: Excretion
2.6.6.13.1: Introduction to Secretion
2.6.6.13.2: Urinary Excretion
2.6.6.13.3: Fecal Excretion
2.6.6.13.4: Exhaled Air

2.6.6.13.5: Other Routes
2.6.6.14: Cellular Toxicology
2.6.6.14.1: Adaptation
2.6.6.14.2: Cell Damage and Tissue Repair
2.6.6.14.3: Cancer
2.6.6.14.4: Neurotoxicity
2.6.6.15: Intuitive Toxicology and Risk Communication
2.6.6.15.1: Intuitive Toxicology
2.6.6.15.2: Risk Communication
2.6.6.16: Environmental Toxicology, Environmental Health, and One Health
2.6.6.16.1: Environmental Toxicology
2.6.6.16.2: Environmental Health
2.6.6.16.3: One Health
2.6.6.17: Conclusion
Detailed Licensing
Index
Glossary
Detailed Licensing

00: Front Matter

This page was auto-generated because a user created a sub-page to this page.

2: Environmental Chemistry

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

This text was compiled on 05/03/2026

TABLE OF CONTENTS

Licensing

1: Introduction

- 1.1: Environmental Science Preface
- 1.2: What Is Science, and How Does It Work?
  - 1.2.1: What is Science?
  - 1.2.2: The Scientific Method
  - 1.2.3: Scientific Papers
  - 1.2.4: Basic and Applied Science
  - 1.2.5: Data Dive- Tonsillectomy Trends
  - 1.2.6: Review
- 1.3: Math Blast- An Overview of Essential Mathematics Used in Science
  - 1.3.1: Getting down the basics- What are the essential mathematical and statistical terms and concepts?
  - 1.3.2: Summarizing the data- Descriptive statistics
  - 1.3.3: Putting it all together- Inferential statistics and hypothesis testing
  - 1.3.4: Data Dive- California Wildfires
  - 1.3.5: Review

2: Environmental Chemistry

- 00: Front Matter
  - TitlePage
  - InfoPage
  - Table of Contents
  - Licensing
- 2.1: Geochemistry (Lower)
  - Licensing
  - 2.1.1: The Earth and its Lithosphere
    - 2.1.1.1: Composition and Structure of the Earth
    - 2.1.1.2: Origin of the Elements
    - 2.1.1.3: Formation and evolution of the Earth
    - 2.1.1.4: The Earth's crust
  - 2.1.2: The Hydrosphere
    - 2.1.2.1: Water, Water Everywhere...
    - 2.1.2.2: The hydrosphere and the oceans
    - 2.1.2.3: Chemistry and geochemistry of the oceans
    - 2.1.2.4: Chemical budgets of oceanic elements
  - 2.1.3: The Atmosphere
    - 2.1.3.1: Structure and Composition of the Atmosphere
    - 2.1.3.2: Origin and Evolution of the Atmosphere
  - 2.1.4: The Biosphere
    - 2.1.4.1: Chemistry and Energetics of the Life Process
    - 2.1.4.2: Biogeochemical Evolution

- 2.1.4.3: Gaia - Bioregulation of the Environment
- 2.2: Supplemental Modules (Environmental Chemistry)
  - 2.2.1: Acid Rain
    - 2.2.1.1: Acid Rain
    - 2.2.1.2: Acid Rain Transport
    - 2.2.1.3: Acid Snow
    - 2.2.1.4: Electricity Generation
    - 2.2.1.5: Playground 1
    - 2.2.1.6: Sources of Nitrogen Oxides
    - 2.2.1.7: Sources of Sulfur Oxides
  - 2.2.2: Aquatic Chemistry
    - 2.2.2.1: 1. Acid-Base Chemistry of Natural Aquatic Systems
    - 2.2.2.2: 2. Carbonate Equilibria in Natural Waters
    - 2.2.2.3: 3. Redox Equilibria in Natural Waters
    - 2.2.2.4: 4. Solids in Contact With Natural Waters
    - 2.2.2.5: Fundamental Characteristics of Water
    - 2.2.2.6: Natural Water
    - 2.2.2.7: Water Biology
    - 2.2.2.8: Water Chemistry
    - 2.2.2.9: Water Physics
    - 2.2.2.10: Water Treatment
  - 2.2.3: Atmospheric Chemistry
    - 2.2.3.1: Atmosphere
    - 2.2.3.2: Carbon Cycle
    - 2.2.3.3: Carbon Oxides
    - 2.2.3.4: Ozone
    - 2.2.3.5: Photosynthesis
  - 2.2.4: Biochemical Cycles
    - 2.2.4.1: Biochemical Cycles
    - 2.2.4.2: Carbon Cycle
    - 2.2.4.3: Nitrogen Cycle
- 2.3: Environmental Toxicology (van Gestel et al.)
  - 00: Front Matter
    - Table of Contents
  - 2.3.1: Environmental Toxicology
    - 2.3.1.1: Environmental toxicology
    - 2.3.1.2: DPSIR
    - 2.3.1.3: Short history
  - 2.3.2: Environmental Chemistry, Chemicals
    - 2.3.2.1: Introduction
    - 2.3.2.2: Pollutants with specific properties
    - 2.3.2.3: Pollutants with specific use
  - 2.3.3: Environmental Chemistry - From Fate to Exposure
    - 2.3.3.1: Environmental compartments
    - 2.3.3.2: Sources of chemicals
    - 2.3.3.3: Pathways and processes determining chemical fate
    - 2.3.3.4: Partitioning and Partitioning Constants

- 2.3.3.5: Metal Speciation
- 2.3.4: Toxicology
  - 2.3.4.1: Toxicokinetics
  - 2.3.4.2: Toxicodynamics and Molecular Interactions
  - 2.3.4.3: Toxicity Testing
  - 2.3.4.4: Increasing ecological realism in toxicity testing
  - 2.3.4.5: Availability and bioavailability
  - 2.3.4.6: Degradation
  - 2.3.4.7: Modelling exposure
  - 2.3.4.8: Section 4.1. Toxicokinetics
  - 2.3.4.9: Section 4.4. Increasing ecological realism in toxicity testing
  - 2.3.4.10: Section 4.2. Toxicodynamics and Molecular Interactions
  - 2.3.4.11: Section 4.3. Toxicity testing
- 2.3.5: Population, Community and Ecosystem Ecotoxicology
  - 2.3.5.1: Introduction- Linking population, community and ecosystem responses
  - 2.3.5.2: Population ecotoxicology in laboratory settings
  - 2.3.5.3: Wildlife population ecotoxicology
  - 2.3.5.4: Trait-based approaches
  - 2.3.5.5: Population models
  - 2.3.5.6: Metapopulations
  - 2.3.5.7: Community ecotoxicology
  - 2.3.5.8: Structure versus function incl. ecosystem services
  - 2.3.5.9: Landscape ecotoxicology
- 2.3.6: Risk Assessment and Regulation
  - 2.3.6.1: Introduction - The Essence of Risk Assessment
  - 2.3.6.2: Ecosystem services and protection goals
  - 2.3.6.3: Predictive risk assessment approaches and tools
  - 2.3.6.4: Diagnostic risk assessment approaches and tools
  - 2.3.6.5: Regulatory Frameworks
  - 2.3.6.6: Risk management and risk communication
  - 2.3.6.7: Risk perception
- 2.4: Key Elements of Green Chemistry (Lucia)
  - Licensing
  - 2.4.1: Key Elements of Green Chemistry
    - 2.4.1.1: Introduction to Green Chemistry
    - 2.4.1.2: Green Chemistry Concepts
    - 2.4.1.3: Case Study
    - 2.4.1.4: Conclusions and Review
  - 2.4.2: Life-Cycle Analysis
    - 2.4.2.1: Introduction to Life Cycles
    - 2.4.2.2: LCA/LCIA Concepts
    - 2.4.2.3: Conclusions and Review
  - 2.4.3: Hazards
    - 2.4.3.1: Introduction to Hazards of Chemistry
    - 2.4.3.2: Hazard Concepts
    - 2.4.3.3: Case Study - Badger Army Ammunition Plant
    - 2.4.3.4: Green Technologies for Safer Chemical Production

- 2.4.3.5: Conclusions and Review Questions
- 2.4.4: Alternative Solvents
  - 2.4.4.1: Introduction to Solvents
  - 2.4.4.2: Solvent Concepts
  - 2.4.4.3: Solvent Alternatives
  - 2.4.4.4: Conclusions and Review
- 2.4.5: Alternative Reagants
  - 2.4.5.1: Introduction to Reagants
  - 2.4.5.2: Reagant Concepts
  - 2.4.5.3: Specific Replacements
  - 2.4.5.4: Conclusions and Review
- 2.4.6: Reaction Types, Design, and Efficiency
  - 2.4.6.1: Introduction to Reactions
  - 2.4.6.2: Reaction Concepts
  - 2.4.6.3: Reaction Design Concepts
  - 2.4.6.4: Conclusions and Review
- Detailed Licensing

- 2.5: Green Chemistry and the Ten Commandments of Sustainability (Manahan)
  - Licensing
  - 2.5.1: Sustainability and the Environment
    - 2.5.1.1: Sustainability
    - 2.5.1.2: The Environment and the Five Environmental Sphere
    - 2.5.1.3: Seeing Green
    - 2.5.1.4: Natural Capital of the Earth
    - 2.5.1.5: Sustainability as a Group Effort- It Takes a (Very Big) Village
    - 2.5.1.6: Sustainable Energy- Away from the Sun and Back Again
    - 2.5.1.7: Green Science
    - 2.5.1.8: Green Technology
    - 2.5.1.9: Sustainability and the Eco-Economy
    - 2.5.1.10: Green Products and Services- Design for Sustainability
    - 2.5.1.11: Questions and Problems
  - 2.5.2: The Key Role of Chemistry and Making Chemistry Green
    - 2.5.2.1: Chemistry is Good (and Unavoidable)
    - 2.5.2.2: The Environment and its Chemistry
    - 2.5.2.3: What is Environmental Chemistry?
    - 2.5.2.4: Environmental Pollution
    - 2.5.2.5: Practice of Green Chemistry
    - 2.5.2.6: Green Chemistry and Synthetic Chemistry
    - 2.5.2.7: Reduction of Risk- Hazard and Exposure
    - 2.5.2.8: The Risks of No Risks
    - 2.5.2.9: Waste Prevention
    - 2.5.2.10: Basic Principles of Green Chemistry
    - 2.5.2.11: Some Things to Know About Chemistry before You Even Start
    - 2.5.2.12: Combining Atoms to Make Molecules and Compounds
    - 2.5.2.13: The process of making and breaking chemical bonds- chemical reactions
    - 2.5.2.14: The Nature of Matter and States of Matter
    - 2.5.2.15: References and Questions

- 2.5.3: The Elements - Basic Building Blocks of Green Chemicals
  - 2.5.3.1: Elements, Atoms, and Atomic Theory
  - 2.5.3.2: Hydrogen - The Simplest Atom
  - 2.5.3.3: Helium - The First Noble Gas
  - 2.5.3.4: Lithium, The First Metal
  - 2.5.3.5: The Second Period of the Periodic Table
  - 2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons
  - 2.5.3.7: Completing the 20-Element Periodic Table
  - 2.5.3.8: The Brief Periodic Table is Complete
  - 2.5.3.E: The Elements - Basic Building Blocks of Green Chemicals (Exercises)
  - 2.5.3.R: References
- 2.5.4: Compounds- Safer Materials for a Safer World
  - 2.5.4.1: Chemical Bonds and Compound Formation
  - 2.5.4.2: Electrons Involved in Chemical Bonds and Octets of Electrons
  - 2.5.4.3: Sodium Chloride and Ionic Bonds
  - 2.5.4.4: Covalent Bonds in H2 and Other Molecules
  - 2.5.4.5: Covalent Bonds in Compounds
  - 2.5.4.6: Covalent Bonds and Green Chemistry
  - 2.5.4.7: Predicting Formulas of Covalently Bound Compounds
  - 2.5.4.8: Chemical Formulas, the Mole, and Percentage Composition
  - 2.5.4.9: What Are Chemical Compounds Called?
  - 2.5.4.10: Acids, Bases and Salts
  - 2.5.4.11: Questions and Problems
- 2.5.5: Chemical Reactions- Making Materials Safely and Sustainable
  - 2.5.5.1: Describing What Happens with Chemical Equations
  - 2.5.5.2: Balancing Chemical Equations
  - 2.5.5.3: Just Because You Can Write It Doesn't Mean That It Will Happen
  - 2.5.5.4: Yield And Atom Economy in Chemical Reactions
  - 2.5.5.5: Catalysts That Make Reactions Go
  - 2.5.5.6: Kinds of Chemical Reactions
  - 2.5.5.7: Oxidation-Reduction Reactions and Green Chemistry
  - 2.5.5.8: Quantitative Information from Chemical Reactions
  - 2.5.5.9: Energy in Chemical Reactions
  - 2.5.5.10: Stoichiometry by the Mole Ratio Method
  - 2.5.5.11: Limiting Reactant and Percent Yield
  - 2.5.5.12: Titrations - Measuring Moles by Volume of Solution
  - 2.5.5.13: Industrial Chemical Reactions - The Solvay Process
  - 2.5.5.14: Questions and Problems
- 2.5.6: The Wonderful World of Carbon - Organic Chemistry and Biochemicals
  - 2.5.6.1: Rings and Chains of Carbon Atoms
  - 2.5.6.2: Compounds of Carbon and Hydrogen - Hydrocarbons
  - 2.5.6.3: Using Lines To Show Organic Structural Formulas
  - 2.5.6.4: Functional Groups
  - 2.5.6.5: Giant Molecules from Small Organic Molecules
  - 2.5.6.6: Questions and Problems
  - 2.5.6.7: Supplementary References
- 2.5.7: Chemistry of Life and Green Chemistry
  - 2.5.7.1: Green Biochemistry

- 2.5.7.2: Biochemistry and the Cells
- 2.5.7.3: Carbohydrates
- 2.5.7.4: Proteins
- 2.5.7.5: Lipids - Fats, Oils, and Hormones
- 2.5.7.6: Nucleic Acids
- 2.5.7.7: Enzymes
- 2.5.7.8: Biochemical Processes in Metabolism
- 2.5.7.9: Biochemistry of Toxic Substances and Toxicological Chemistry
- 2.5.7.10: Literature Cited
- 2.5.7.11: Questions and Problems
- 2.5.7.12: Supplementary References

- 2.5.8: The Five Environmental Spheres and Biogeochemical Cycle
  - 2.5.8.1: Five Environmental Spheres
  - 2.5.8.2: The Hydrosphere
  - 2.5.8.3: The Atmosphere
  - 2.5.8.4: The Geosphere
  - 2.5.8.5: The Biosphere
  - 2.5.8.6: The Anthrosphere
  - 2.5.8.7: Cycles of Matter
  - 2.5.8.8: Questions and Problems
  - 2.5.8.9: Supplementary References
- 2.5.9: Water - the Ultimate Green Substance
  - 2.5.9.1: H_2O - Simple Formula, Remarkable Molecule
  - 2.5.9.2: Occurrence, Availability , and Utilization of Water on Earth
  - 2.5.9.3: Water Chemistry
  - 2.5.9.4: Water Pollution
  - 2.5.9.5: Greening of Water - Purification Before Use
  - 2.5.9.6: Wastewater Treatment
  - 2.5.9.7: Advanced Wastewater Treatment and Recycling
  - 2.5.9.8: The Many Uses of Water
  - 2.5.9.9: Hot Water - Pressurized Subcritical Water
  - 2.5.9.10: Supercritical Water
  - 2.5.9.11: Literature Cited
  - 2.5.9.12: Questions and Problems
  - 2.5.9.13: Supplementary References
- 2.5.10: Blue Skies for a Green Environment
  - 2.5.10.1: More Than Just Air to Breathe
  - 2.5.10.2: The Gas Laws
  - 2.5.10.3: The Protective Atmosphere
  - 2.5.10.4: Atmospheric Chemistry and Photochemical Reactions
  - 2.5.10.5: The Enormous Importance of Climate
  - 2.5.10.6: Atmospheric Particle Pollutants
  - 2.5.10.7: Pollutant Gaseous Oxides
  - 2.5.10.8: Acid Rain
  - 2.5.10.9: Miscellaneous Gases in the Atmosphere
  - 2.5.10.10: Photochemical Smog
  - 2.5.10.11: Natural Capital of the Atmosphere
  - 2.5.10.12: Energy and Mass Transfer in the Atmosphere
  - 2.5.10.13: Questions and Problems

- 2.5.10.14: Supplementary References
- 2.5.11: The Geosphere and a Green Earth
  - 2.5.11.1: Probing Deep into the Geosphere - A Well Too Far?
  - 2.5.11.2: The Nature of the Geosphere
  - 2.5.11.3: The Geosphere as a Source of Natural Capital
  - 2.5.11.4: Environmental Hazards of the Geosphere
  - 2.5.11.5: Water on and in the Geosphere
  - 2.5.11.6: Anthrospheric Influences on the Geosphere
  - 2.5.11.7: The Geosphere as a Waste Repository
  - 2.5.11.8: Have You Thanked a Clod Today?
  - 2.5.11.9: Plant Nutrients and Fertilizers in Soil
  - 2.5.11.10: Soil and Plants Related To Wastes and Pollutants
  - 2.5.11.11: Soil Loss - Desertification and Deforestation
  - 2.5.11.12: Literature Cited and Supplementary References
  - 2.5.11.13: Questions and Problems
- 2.5.12: The Biosphere and the Role of Green Chemistry in Feeding a Hungry World
  - 2.5.12.1: Pigweed's Revenge
  - 2.5.12.2: The Biosphere
  - 2.5.12.3: Cells - Basic Units of Life
  - 2.5.12.4: Metabolism and Control in Organisms
  - 2.5.12.5: Reproduction and Inherited Traits
  - 2.5.12.6: Stability and Equilibrium of the Biosphere
  - 2.5.12.7: DNA and the Human Genome
  - 2.5.12.8: Genetic Engineering
  - 2.5.12.9: Biological Interaction with Environmental Chemicals
  - 2.5.12.10: Biodegradation
  - 2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture
  - 2.5.12.12: Agricultural Applications of Genetically Modified Organisms
  - 2.5.12.13: The Anthrosphere in Support of the Biosphere
  - 2.5.12.14: Livestock and their Wastes
  - 2.5.12.15: Questions and Problems
  - 2.5.12.16: Supplementary References
- 2.5.13: The Anthrosphere, Industrial Ecology, and Green Chemistry
  - 2.5.13.1: Industrial Ecology and Industrial Ecosystems
  - 2.5.13.2: Metabolic Processes in Industrial Ecosystems
  - 2.5.13.3: Life Cycles in Industrial Ecosystems
  - 2.5.13.4: Kinds of Products
  - 2.5.13.5: Attributes Required by an Industrial Ecosystem
  - 2.5.13.6: The Kalundborg Industrial Ecosystem
  - 2.5.13.7: Environmental Impacts of Industrial Ecosystems
  - 2.5.13.8: Green Chemistry and Industrial Ecology
  - 2.5.13.9: Predicting and Reducing Hazards
  - 2.5.13.10: The E-Factor in Green Chemistry
  - 2.5.13.11: Catalysts and Catalysis
  - 2.5.13.12: Biocatalysis with Enzymes
  - 2.5.13.13: Energizing Chemical Reactions and Process Intensification
  - 2.5.13.14: Solvents and Alternate Reaction Media
  - 2.5.13.15: Feedstocks and Reagents
  - 2.5.13.16: Literature Cited and Supplementary References

- 2.5.13.17: Questions and Problems
- 2.5.14: Feeding the Anthrosphere- Utilizing Renewable and Biological Materials
  - 2.5.14.1: Feeding the Anthrosphere
  - 2.5.14.2: Utilization of Feedstocks
  - 2.5.14.3: Biological Feedstocks
  - 2.5.14.4: Biological Sources of Chemicals
  - 2.5.14.5: Biorefineries and Biomass Utilization
  - 2.5.14.6: Monosaccharide Feedstocks - Glucose, Fructose, and Xylose
  - 2.5.14.7: Cellulose
  - 2.5.14.8: Feedstocks from Cellulose Wastes
  - 2.5.14.9: Lignin
  - 2.5.14.10: Direct Biosynthesis of Polymers
  - 2.5.14.11: Bioconversion Processes for Synthetic Chemicals
  - 2.5.14.12: Bamboo - Ancient Material for the Future
  - 2.5.14.13: Literature Cited
  - 2.5.14.14: Questions and Problems
- 2.5.15: Sustainable Energy- The Essential Basis of Green Systems
  - 2.5.15.1: Sustainability and Energy
  - 2.5.15.2: What is Energy?
  - 2.5.15.3: Radiant Energy from the Sun
  - 2.5.15.4: Sources of Energy Used in the Anthrosphere
  - 2.5.15.5: Conversions Between Forms of Energy
  - 2.5.15.6: Green Technology for Energy Conversion
  - 2.5.15.7: Energy Conservation
  - 2.5.15.8: Depletable Fossil Fuels
  - 2.5.15.9: Carbon Sequestration for Fossil Fuel Utilization
  - 2.5.15.10: Nuclear Energy
  - 2.5.15.11: Renewable Energy Sources - Solar Energy
  - 2.5.15.12: Energy from Wind and Water
  - 2.5.15.13: Biomass Energy
  - 2.5.15.14: Geothermal Energy
  - 2.5.15.15: Hydrogen for Energy Storage and Utilization
  - 2.5.15.16: Literature Cited and Supplementary References
  - 2.5.15.17: Questions and Problems
- 2.5.16: Terrorism, Toxicity, and Vulnerability- Green Chemistry and Technology in Defense of Human Welfare
  - 2.5.16.1: Vulnerability to Terrorist Attack
  - 2.5.16.2: Protecting the Anthrosphere
  - 2.5.16.3: Substances That Explode, Burn, or React Violently
  - 2.5.16.4: Toxic Substances and Toxicology
  - 2.5.16.5: Toxic Chemical Attack
  - 2.5.16.6: Protecting Water, Food, and Air
  - 2.5.16.7: Detecting Hazards
  - 2.5.16.8: Green Chemistry to Combat Terrorism
  - 2.5.16.9: Green Chemistry for Sustainable Prosperity and a Safer World
  - 2.5.16.10: Green Chemistry to Combat Terrorism
  - 2.5.16.11: Green Chemistry for Sustainable Prosperity and and a Safer World
  - 2.5.16.12: Literature Cited
  - 2.5.16.13: Questions and Problems
- 2.5.17: The Ten Commandments of Sustainability and Sensible Measure

- 2.5.17.1: We Cannot Go on Like This
- 2.5.17.2: The First Commandment
- 2.5.17.3: The Second Commandment
- 2.5.17.4: The Third Commandment
- 2.5.17.5: The Fourth Commandment
- 2.5.17.6: The Fifth Commandment
- 2.5.17.7: The Sixth Commandment
- 2.5.17.8: The Seventh Commandment
- 2.5.17.9: The Ninth Commandment
- 2.5.17.10: The Tenth Commandment
- 2.5.17.11: Some Sensible Measures for Sustainability
- 2.5.17.12: The Eighth Commandment
- 2.5.17.13: Literature Cited
- 2.5.17.14: Questions and Problems
- 2.5.17.15: Supplementary References
- Detailed Licensing

- 2.6: Toxicology MSDT
  - Licensing
  - Preface
  - 2.6.1: Pathophysiology
    - 2.6.1.1: What is Pathophysiology?
    - 2.6.1.2: Outcomes of Targeted and Non-Targeted Toxicity
    - 2.6.1.3: Cellular Response to Toxicant-Induced Injury
    - 2.6.1.4: Repair and Adaptation
    - 2.6.1.5: Patterns of Toxic Injury
    - 2.6.1.6: Final Evaluation
  - 2.6.2: Biochemistry and Molecular Genetics
    - 2.6.2.1: Introduction to Biomolecules and Cell Components
    - 2.6.2.2: Cell Structure and Subcellular Compartments
    - 2.6.2.3: DNA and RNA Metabolism
    - 2.6.2.4: Epigenetic Mechanisms
    - 2.6.2.5: Section 2 Final Evaluation
  - 2.6.3: Principles of Genetic Toxicology
    - 2.6.3.1: Introduction to Genetic-toxicology Assay
    - 2.6.3.2: Different Genetic Damages or Mutations
    - 2.6.3.3: Different Genetic-Toxicology Assays
    - 2.6.3.4: Different Cytotoxicity Assays
    - 2.6.3.5: Epigenetics Assay
    - 2.6.3.6: Section 3 Final Evaluation
  - 2.6.4: Applied Systems Toxicology
    - 2.6.4.1: Systems Toxicology
    - 2.6.4.2: Dose Level and Applied Toxicology
    - 2.6.4.3: Tools and Technologies in Systems Toxicology
    - 2.6.4.4: Other Approaches for Predictive Toxicity Modeling
    - 2.6.4.5: Technologies Used In Systems Biology/Toxicology
    - 2.6.4.6: Takeaways Summary
    - 2.6.4.7: Section 4 Final Evaluation

- 2.6.5: Regulatory Toxicology
  - 2.6.5.1: Introduction to Regulatory Toxicology
  - 2.6.5.2: Global Regulatory Toxicology
  - 2.6.5.3: Topic 3- Regional Regulatory Toxicology
  - 2.6.5.4: National Regulatory Toxicology
  - 2.6.5.5: State Regulatory Toxicology
  - 2.6.5.6: Non-Governmental Regulatory Toxicology
  - 2.6.5.7: Section 5 Final Evaluation
- 2.6.6: Principles of Toxicology
  - 2.6.6.1: Introduction to Toxicology
    - 2.6.6.1.1: What is toxicology?
    - 2.6.6.1.2: Basic Terminology
  - 2.6.6.2: Dose and Dose Response
    - 2.6.6.2.1: Dose and It's Impact on Toxicity
    - 2.6.6.2.2: The Dose Response Relationship
    - 2.6.6.2.3: Dose Estimates of Toxic Effects
    - 2.6.6.2.4: Determining the Safety of a Drug
    - 2.6.6.2.5: NOAEL and LOAEL
  - 2.6.6.3: Toxic Effects
    - 2.6.6.3.1: Types of Toxic Effects
    - 2.6.6.3.2: Factors Affecting Toxicity
    - 2.6.6.3.3: Systemic Toxic Effects
    - 2.6.6.3.4: Organ Specific Toxic Effects
  - 2.6.6.4: Interactions
    - 2.6.6.4.1: Interactions
  - 2.6.6.5: Toxicity Testing Methods
    - 2.6.6.5.1: Testing and Assessing Toxicity
    - 2.6.6.5.2: Clinical Investigations and Other Types of Human Data
    - 2.6.6.5.3: Epidemiology Studies
  - 2.6.6.6: Risk Assessment
    - 2.6.6.6.1: Risk Assessment
    - 2.6.6.6.2: Hazard Identification
    - 2.6.6.6.3: Dose-Response Assessment
    - 2.6.6.6.4: Exposure Assessment
    - 2.6.6.6.5: Risk Characterization
  - 2.6.6.7: Exposure Standards and Guidelines
    - 2.6.6.7.1: Exposure Standards and Guidelines
    - 2.6.6.7.2: Regulation of Consumer Products and Drug Safety
    - 2.6.6.7.3: Environmental Exposure Standards/Guidelines
    - 2.6.6.7.4: Occupational (Workplace) Exposure Standards/Guidelines/Approaches
  - 2.6.6.8: Basic Physiology
    - 2.6.6.8.1: Introduction to Basic Physiology
    - 2.6.6.8.2: Homeostasis
    - 2.6.6.8.3: Organs and Organ Systems
    - 2.6.6.8.4: Tissues
    - 2.6.6.8.5: Cells

- 2.6.6.8.6: Chemicals
- 2.6.6.9: Introduction to Toxicokinetics
  - 2.6.6.9.1: What is Toxicokinetics
- 2.6.6.10: Absorption
  - 2.6.6.10.1: Introduction to Absorption
  - 2.6.6.10.2: Gastrointestinal Tract
  - 2.6.6.10.3: Respiratory Tract
  - 2.6.6.10.4: Dermal Route
  - 2.6.6.10.5: Other Routes of Exposure
- 2.6.6.11: Distribution
  - 2.6.6.11.1: Introduction to Distribution
  - 2.6.6.11.2: Influence of Route of Exposure
  - 2.6.6.11.3: Disposition Models
  - 2.6.6.11.4: Structural Barriers to Distribution
  - 2.6.6.11.5: Storage Sites
- 2.6.6.12: Biotransformation
  - 2.6.6.12.1: Introduction to Biotransformation
  - 2.6.6.12.2: Chemical Reactions
  - 2.6.6.12.3: Biotransformation Sites
  - 2.6.6.12.4: Modifiers of Biotransformation
- 2.6.6.13: Excretion
  - 2.6.6.13.1: Introduction to Secretion
  - 2.6.6.13.2: Urinary Excretion
  - 2.6.6.13.3: Fecal Excretion
  - 2.6.6.13.4: Exhaled Air
  - 2.6.6.13.5: Other Routes
- 2.6.6.14: Cellular Toxicology
  - 2.6.6.14.1: Adaptation
  - 2.6.6.14.2: Cell Damage and Tissue Repair
  - 2.6.6.14.3: Cancer
  - 2.6.6.14.4: Neurotoxicity
- 2.6.6.15: Intuitive Toxicology and Risk Communication
  - 2.6.6.15.1: Intuitive Toxicology
  - 2.6.6.15.2: Risk Communication
- 2.6.6.16: Environmental Toxicology, Environmental Health, and One Health
  - 2.6.6.16.1: Environmental Toxicology
  - 2.6.6.16.2: Environmental Health
  - 2.6.6.16.3: One Health
- 2.6.6.17: Conclusion
- Detailed Licensing

- Index
- Glossary
- Detailed Licensing

3: Ecology

- 3.1: Ecology Overview
- 3.2: Populations
  - 3.2.1: Population Dispersion
  - 3.2.2: Population Size
  - 3.2.3: Population Growth and Regulation
  - 3.2.4: Life History
  - 3.2.5: Data Dive- Wildlife Corridors
  - 3.2.6: Review
- 3.3: Communities
  - 3.3.1: Biotic Interactions
    - 3.3.1.1: Trophic Interactions
      - 3.3.1.1.1: Predation
      - 3.3.1.1.2: Herbivory
      - 3.3.1.1.3: Parasitism
      - 3.3.1.1.4: Food Chains and Food Webs
    - 3.3.1.2: Competition
    - 3.3.1.3: Facilitation
    - 3.3.1.4: Symbiosis
  - 3.3.2: Community Structure and Dynamics
  - 3.3.3: Data Dive- Bison Impacts In Prairies
  - 3.3.4: Review
- 3.4: Ecosystems
  - 3.4.1: Ecosystem Types and Dynamics
  - 3.4.2: Matter
  - 3.4.3: Biogeochemical Cycles
  - 3.4.4: Soils
  - 3.4.5: Soil Degradation
  - 3.4.6: Data Dive- Beaver Impacts on Wetlands
  - 3.4.7: Review
- 3.5: Biomes
  - 3.5.1: Climate and Biomes
  - 3.5.2: Terrestrial Biomes
  - 3.5.3: Aquatic Biomes
  - 3.5.4: Data Dive- Biome Carbon Storage
  - 3.5.5: Review

4: Conservation

- 4.1: The Value of Biodiversity
  - 4.1.1: Ecosystem Diversity
  - 4.1.2: Species Diversity
  - 4.1.3: Genetic Diversity
  - 4.1.4: Patterns of Biodiversity
  - 4.1.5: Data Dive- Biodiversity and Drugs
  - 4.1.6: Review
- 4.2: Threats to Biodiversity

- 4.2.1: Extinction
- 4.2.2: Measures of Biodiversity Loss
- 4.2.3: Habitat Loss
- 4.2.4: Overexploitation
- 4.2.5: Pollution
- 4.2.6: Invasive Species
- 4.2.7: Climate Change
- 4.2.8: Data Dive- Threats to Biodiversity
- 4.2.9: Review
- 4.3: Protecting Biodiveristy
  - 4.3.1: Policies
  - 4.3.2: Non-profit Organizations
  - 4.3.3: Species-level Conservation
  - 4.3.4: Protected Areas
  - 4.3.5: Ecosystem Restoration
  - 4.3.6: Economic Influences on Conservation
  - 4.3.7: Individual Choices
  - 4.3.8: Data Dive- Island Fox Populations
  - 4.3.9: Review

5: Humans and the Environment

- 5.1: The Human Population
  - 5.1.1: History of Human Population Growth
  - 5.1.2: The Rate of Human Population Growth
  - 5.1.3: Looking Ahead
  - 5.1.4: Data Dive- World Population Densities
  - 5.1.5: Review
- 5.2: Water Resources
  - 5.2.1: Fresh Water Supply and the Water Cycle
  - 5.2.2: Water Usage
  - 5.2.3: Water Scarcity and Solutions
  - 5.2.4: Data Dive- Aqueducts of Rome
  - 5.2.5: Review
- 5.3: Agriculture
  - 5.3.1: Food Security
  - 5.3.2: Industrial Agriculture
  - 5.3.3: Selective Breeding and Genetic Engineering
  - 5.3.4: Sustainable Agriculture
  - 5.3.5: Data Dive- Landslides and Ag Terraces
  - 5.3.6: Review
- 5.4: Environmental Health
  - 5.4.1: Types of Environmental Hazards
  - 5.4.2: Epidemiology
  - 5.4.3: Infectious Diseases
  - 5.4.4: Environmental Toxicology
  - 5.4.5: Environmental Hazard Reduction
  - 5.4.6: Data Dive- Cholera Cases Worldwide
  - 5.4.7: Review

6: Energy

- 6.1: Fossil Fuels
  - 6.1.1: Types of Fossil Fuels and Formation
  - 6.1.2: Mining, Processing, and Generating Electricity
  - 6.1.3: Fossil Fuel Consumption
  - 6.1.4: Consequences of Fossil Fuels
  - 6.1.5: Data Dive- Global Fossil Fuel Consumption
  - 6.1.6: Review
- 6.2: Nuclear Energy
  - 6.2.1: Radioactive Isotopes
  - 6.2.2: Generating Electricity with Nuclear Energy
  - 6.2.3: Nuclear Energy Consumption
  - 6.2.4: Consequences of Nuclear Energy
  - 6.2.5: Data Dive- Global Nuclear Power Generation
  - 6.2.6: Review
- 6.3: Renewable Energy
  - 6.3.1: Renewable Energy History and Consumption
  - 6.3.2: Wind Energy
  - 6.3.3: Solar Energy
  - 6.3.4: Geothermal Energy
  - 6.3.5: Hydropower
  - 6.3.6: Biofuels (Biomass Energy)
  - 6.3.7: Energy Conservation
  - 6.3.8: Data Dive- Global Renewable Energy Generation
  - 6.3.9: Review

7: Environmental Impacts

- 7.1: Solid Waste Management
  - 7.1.1: Waste Generation
  - 7.1.2: Waste Disposal
  - 7.1.3: Solid Waste and Marine Life
  - 7.1.4: Waste Reduction
  - 7.1.5: Data Dive- "Flushable" Wipes
  - 7.1.6: Review
- 7.2: Pollution
  - 7.2.1: Water Pollution
    - 7.2.1.1: Water Pollutants and Their Sources
    - 7.2.1.2: Water Treatment
    - 7.2.1.3: Mitigating Water Pollution
  - 7.2.2: Air Pollution
    - 7.2.2.1: Ozone Depletion
    - 7.2.2.2: Acid Deposition
  - 7.2.3: Data Dive- Plastic in World Oceans
  - 7.2.4: Review
- 7.3: Climate Change
  - 7.3.1: The Greenhouse Effect and Climate Change

- 7.3.2: Data Dive- Flooding in Venice Italy
- 7.3.3: Review

8: Sustainability

- 8.1: Introduction to Sustainability
- 8.2: History of Sustainability
- 8.3: Environmental Ethics
  - 8.3.1: Environmental Justice and Indigenous Struggles
  - 8.3.2: Perspectives of the Environment
- 8.4: Environmental Policy
- 8.5: Sustainable Cities
- 8.6: Sustainability and the Future
- 8.7: Data Dive- Sustainable Investments
- 8.8: Review

Index

Glossary

Detailed Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

SECTION OVERVIEW

2.1: Geochemistry (Lower)

Geochemistry is the study of the composition and chemical changes that take place on the Earth and within its major divisions: the *lithosphere*, *atmosphere*, *hydrosphere*, and *biosphere*. Environmental chemistry focuses more specifically on the transport of substances within and between these *geospheres* and their subdivisions, with special emphasis on the effects brought about by the activities of humans. The realm of geochemistry extends beyond the Earth, encompassing the entire Solar System and has made important contributions to the understanding of a number of processes including mantle convection, the formation of planets and the origins of granite and basalt.

Licensing

2.1.1: The Earth and its Lithosphere

- 2.1.1.1: Composition and Structure of the Earth
- 2.1.1.2: Origin of the Elements
- 2.1.1.3: Formation and evolution of the Earth
- 2.1.1.4: The Earth's crust

2.1.2: The Hydrosphere

- 2.1.2.1: Water, Water Everywhere...
- 2.1.2.2: The hydrosphere and the oceans
- 2.1.2.3: Chemistry and geochemistry of the oceans
- 2.1.2.4: Chemical budgets of oceanic elements

2.1.3: The Atmosphere

- 2.1.3.1: Structure and Composition of the Atmosphere
- 2.1.3.2: Origin and Evolution of the Atmosphere

2.1.4: The Biosphere

- 2.1.4.1: Chemistry and Energetics of the Life Process
- 2.1.4.2: Biogeochemical Evolution
- 2.1.4.3: Gaia - Bioregulation of the Environment

2.1: Geochemistry (Lower) is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

Front Matter

Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

SECTION OVERVIEW

2.1.1: The Earth and its Lithosphere

2.1.1.1: Composition and Structure of the Earth

2.1.1.2: Origin of the Elements

2.1.1.3: Formation and evolution of the Earth

2.1.1.4: The Earth's crust

This page titled 2.1.1: The Earth and its Lithosphere is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

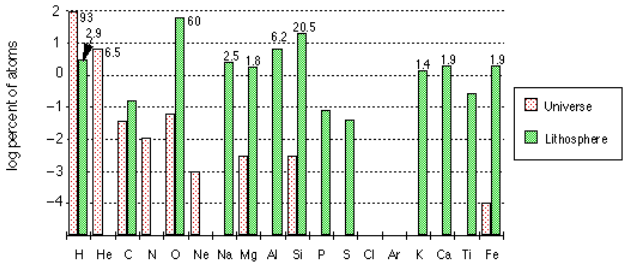
2.1.1.1: Composition and Structure of the Earth

The earth has been in a state of continual change since its formation. The major part of this change, involving volcanism and tectonics, has been driven by heat produced from the decay of radioactive elements within the earth. The other source of change has been solar energy, which acts as the driving force of weathering and is the ultimate source of energy for living organisms.

The solar system was probably formed about 4.6 billion years ago, and the oldest known rocks have an age of 3.8 billion years. There is thus a gap of 0.8 billion years for which there is no direct evidence. It is known that the earth was subjected to extensive bombardment earlier in its history; recent computer simulations suggest that the moon could have resulted from an especially massive collision with another body. Although these major collisions have diminished in magnitude as the matter in the solar system has become more consolidated, they continue to occur, with the most recent one being responsible for the annihilation of the dinosaurs and much of the other life on Earth. The lack of many overt signs of these collisions (such as craters, for example) testifies to the dynamic processes at work on the Earth's surface and beneath it.

Chemical Composition of the Earth

The earth is composed of 90 chemical elements, of which 81 have at least one stable isotope. The unstable elements are ^{43}Tc and ^{61}Pm, and all elements heavier than ^{83}Bi.



> *Note that the vertical axis is logarithmic, which has the effect of greatly reducing the visual impression of the differences between the various elements.*

The chart gives the abundances of the elements present in the solar system, in the earth as a whole, and in the various geospheres. Of particular interest are the differences between the terrestrial and cosmic abundances, which are especially notable in the cases of the lighter elements (H, C, N) and the noble gas elements (He, Ne, Ar, Xe, Kr).

Example 2.1.1.1.1

Given the mix of elements that are present in the earth, how might they combine so as to produce the chemical composition we now observe?

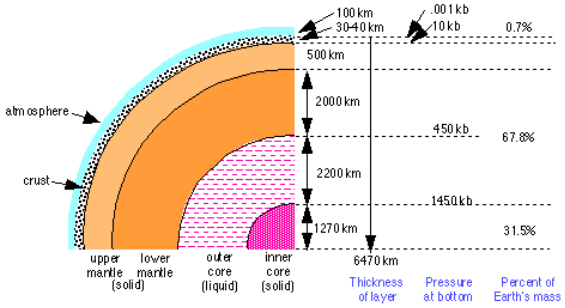
Solution

Given the mix of elements that are present in the earth, how might they combine so as to produce the chemical composition we now observe? Thermodynamics allows us to predict the composition that any isolated system will eventually reach at a given temperature and pressure. Of course the earth is not an isolated system, although most parts of it can be considered approximately so in many respects, on time scales sufficient to make thermodynamic predictions reasonably meaningful. The equilibrium states predicted by thermodynamics differ markedly from the observed compositions. The atmosphere, for example, contains 0.03% CO_2, 78% N_2 and 21% O_2; in a world at equilibrium the air would be 99% CO_2.

Similarly, the oceans, containing about 3.5% NaCl, would have a salt content of 35% if they were in equilibrium with the atmosphere and the lithosphere. Trying to understand the mechanisms that maintain these non-equilibrium states is an important part of contemporary environmental geochemistry.

Structure of the Earth

Studies based on the reflection and refraction of the acoustic waves resulting from earthquakes show that the interior of the earth consists of four distinct regions. A combination of physical and chemical processes led to the differentiation of the earth into these major parts. This is believed to have occurred approximately 4 billion years ago.



The Earth's Core

The Earth’s core is believed to consist of two regions. The inner core is solid, while the outer core is liquid. This phase difference probably reflects a difference in pressure and composition, rather than one of temperature. Density estimates obtained from seismological studies indicate that the core is metallic, and mainly iron, with 8-10 percent of lighter elements.

Hypotheses about the nature of the core must be consistent with the the core’s role as the source of the earth’s magnetic field. This field arises from convective motion of the electrically conductive liquid comprising the outer core. Whether this convection is driven by differences in temperature or composition is not certain. The estimated abundance of radioactive isotopes (mainly U^{238} and K^{40} in the core is sufficient to provide the thermal energy required to drive the convective dynamo. Laboratory experiments on the high-pressure behavior of iron oxides and sulfides indicate that these substances are probably metallic in nature, and hence conductive, at the temperatures (4000-5000K) and pressures (1.3-3.5 million atm) that are estimated for the core. Their presence in the core, alloyed with the iron, would be consistent with the observed density, and would also resolve the apparent lack of sulfur in the earth, compared to its primordial abundance.

The mantle

The region extending from the outer part of the core to the crust of the earth is known as the *mantle*. The mantle is composed of oxides and silicates, i.e., of rock. It was once believed that this rock was molten, and served as a source of volcanic magma. It is now known on the basis of seismological evidence that the mantle is not in the liquid state. Laboratory experiments have shown, however, that when rock is subjected to the high temperatures and pressures believed to exist in the mantle, it can be deformed and flows very much like a liquid.

The upper part of the mantle consists of a region of convective cells whose motion is driven by the heat due to decay of radioactive potassium, thorium, and uranium, which were selectively incorporated in the crystal lattices of the lower-density minerals that form the mantle. There are several independent sources of evidence of this motion. First, there are gravitational anomalies; the force of gravity, measured by changes in elevation in the sea surface, is different over upward and downward moving regions, and has permitted the mapping of some of the convective cells. Secondly, numerous isotopic ratio studies have traced the exchange of material between oceanic sediments, upper mantle rock, and back into the continental crust, which forms from melting of the upper mantle. Thirdly, the composition of the basalt formed by upper mantle melting is quite uniform everywhere, suggesting complete mixing of diverse materials incorporated into the mantle over periods of 100 million years.

High-pressure studies in the laboratory have revealed that olivine, a highly abundant substance in the mantle composed of Fe, Mg, Si, and O (and also the principal constituent of meteorites) can undergo a reversible phase change between two forms differing in density. Estimates of conditions within the upper mantle suggest that the this phase change could occur within this region in such as way as to contribute to convection. The most apparent effect of mantle convection is the motion it imparts to the earth’s crust, as evidenced by the the external topography of the earth.

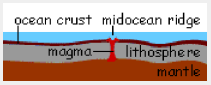
The crust

The outermost part of the earth, known also as the lithosphere, is broken up into plates that are supported by the underlying mantle, and are moved by the convective cells within the mantle at a rate of a few centimetres per year. New crust is formed where plates move away from each other under the oceans, and old crust is recycled back into the mantle as where plates moving in opposite directions collide.

The oceanic crust

The parts of the crust that contain the world's oceans are very different from the parts that form the continents. The continental crust is 10-70 km thick, while oceanic crust averages only 5-7 km in thickness. Oceanic crust is more dense (3.0-3.1 g cm^{-3}) and therefore "floats" on the mantle at a greater depth than does continental crust (density 2.7-2.8). Finally, oceanic crust is much younger; the oldest oceanic crust is about 200 million years old, while the most ancient continental rocks were formed 3.8 billion years ago.

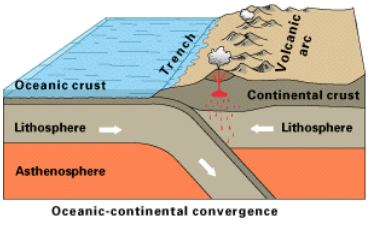
New crust is formed from molten material in the upper mantle at the divergent boundaries that exist at undersea ridges. The melting is due to the rise in temperature associated with the nearly adiabatic decompression of the upper 50-70 km of mantle material as separation of the plates reduces the pressure below. The molten material collects in a magma pocket which is gradually exuded in undersea lava flows. The solidified lava is transformed into crust by the effects of heat and the action of seawater which selectively dissolves the more soluble components.



An animated view of seafloor spreading.

Plate collisions

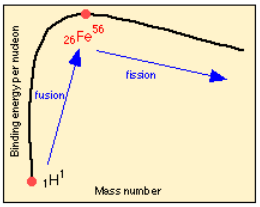
Where two plates collide, one generally plunges under the other and returns to the mantle in a process known as **subduction**. Since the continental plates have a lower density, they tend to float above the oceanic plates and resist subduction. At continental boundaries such as that of the North American west coast where an oceanic plate pushes under the continental crust, oceanic sediments may be sheared off, resulting in a low coastal mountain range (see here for a nice animation of this process.) Also, the injection of water into the subducting material lowers its melting point, resulting in the formation of shallow magma pockets and volcanic activity. Divergent plate boundaries can cross continents, however; temporary divergences create rift valleys such as the Rhine and Rio Grande, while permanent ones eventually lead to new oceanic basins.



Oceanic-continental convergence

Collision of two continental plates can also occur; the most notable example is the one resulting in the formation of the Himalayan mountain chain.

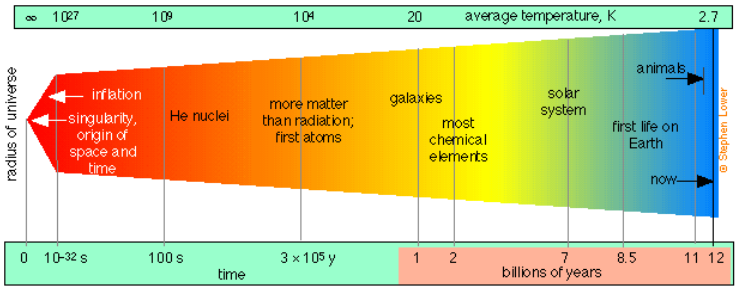
The Earth is composed of 90 chemical elements, of which 81 have at least one stable isotope. Most of these elements have also been detected in stars. Where did these elements come from? The accepted scenario is that the first major element to condense out of the primordial soup was helium , which still comprises about one-quarter of the mass of the known universe.



Hydrogen is the least thermodynamically stable of the elements, and at very high temperatures will combine with itself in a reaction known as *nuclear fusion* to form the next element, $_2He^4$. "Heavier" nuclei (that is , those having high atomic numbers, indicated here by the subscript preceding the element symbol), are more stable than "lighter" ones, so this fusion process can continue up to $_{56}Fe$, which is the most energetically stable of all the nuclides. Beyond this point, heavier nuclei slowly become less stable, so fission becomes more likely. FIssion, however, is not considered an important mechanism of primordial*nucleosynthesis*, so other processes are invoked, as discussed farther below.

Primordial Chemistry

According to the *"big bang" theory* for which there is now overwhelming evidence, the universe as we know it (that is, all space, time, and matter) had its origin in a point source or *singularity* that began an explosive expansion about 12-15 billion years ago, and which is still continuing.



Following a brief period of extremely rapid expansion called *inflation*, protons and neutrons condensed out of the initial quantum soup after about 10^{-32} s. Helium and hydrogen became stable during the first few minutes, along with some of the very lightest nuclides up to $_7Li$, which were formed through various fusion and neutron-absorption processes. Formation of most heavier elements was delayed for about 10^6 years until nucleosynthesis commenced in the first stars. Hydrogen still accounts for about 93% of the atoms in the universe.

The main lines of observational evidence that support this theory are the 2.7K background radiation that permeates the cosmos (the cooled-down remnants of the initial explosion), and the abundances of the lightest elements. Conventional physics is able to extrapolate back to about the first 10^{-33} second; what happened before then remains speculative.

Stellar nucleosynthesis

All elements beyond hydrogen were formed in regions where the concentration of matter was large, and the temperature was high; in other words, in stars. The formation of a star begins when the gravitational forces due to a large local concentration of hydrogen bring about a contraction and compression to densities of around 10^5 g cm^{-3}. This is a highly exothermic process in which the gravitational potential energy is released as heat, about 1200 kJ per gram, raising the temperature to about 10^7 K. Under these conditions, hydrogen nuclei possess sufficient kinetic energy to overcome their electrostatic repulsion and undergo nuclear fusion:

$$4\ {}_1H^1 \rightarrow {}_2\ He^4 + 2b^+ + 2g + 2n \qquad (2.1.1.1.1)$$

Hydrogen burning

There will be a net mass loss in above process, which will therefore be highly exothermic and is known as "hydrogen burning". As hydrogen burning proceeds, the helium collects in the core of the star, raising the density to 10^8 g cm^{-3} and the temperature to 10^8 K. This temperature is high enough to initiate *helium burning*, which proceeds in several steps:

$2\ {}_2He^4 \rightarrow {}_4Be^8 + g$

The first product, ${}_4Be^8$ has a half life of only 10^{-16} sec, but a sufficient amount accumulates to drive the following two reactions:

$$ {}_4Be^8 + {}_2He^4 \rightarrow {}_6C^{12} + g $$

$$ {}_6C^{12} + {}_1H^1 \rightarrow {}_7N^{13} \rightarrow {}_6C^{13} + b^+ + g $$

The size of a star depends on the balance between the kinetic energy of its matter and the gravitational attraction of its mass. As the helium burning runs its course, the temperature drops and the star begins to contract. The course of further nucleosynthesis and the subsequent fate of the star itself depends on the star's mass.

Small stars

If the mass of the star is no greater than 1.4 times the mass of our sun, the star collapses to a white dwarf, and eventually cools to a dark, dense dead star.

Big stars

In larger stars, the gravitational forces are sufficiently strong to overcome the normal repulsion between atoms, and so gravitational collapse continues. The gravitational energy released in this process produces temperatures of 6×10^8 K, which are sufficient to initiate a complex series of nuclear reactions known as the *carbon-nitrogen cycle*. The net reaction of this cycle is the further fusion of hydrogen to helium, in which C^{12} acts as a catalyst, and various nuclides of nitrogen and oxygen are intermediates. The temperature is sufficiently high, however, to initiate fusion reactions of some of these intermediates:

$$ {}_6C^{12} + {}_6C^{12} \rightarrow {}_{10}Ne^{20} + {}_2He^4 $$

$$ 2\ {}_8O^{16} \rightarrow {}_{14}Si^{28} + {}_2He^4 $$

$$ 2\ {}_8O^{16} \rightarrow {}_{16}S^{31} + {}_0n^1 $$

Supernovas

The intense gamma radiation that is produced in some of these reactions breaks some of the product nuclei into smaller fragments, which can then fuse into a variety of heavier species, up to the limit of ${}_{26}Fe^{56}$, beyond which fusion is no longer exothermic. The greater relative abundance of elements such as ${}_6C^{12}$, ${}_8O^{16}$, and ${}_{10}Ne^{20}$ which differ by a ${}_2He^4$ nucleus, reflects the participation of the latter species in these processes. These exothermic reactions eventually produce temperatures of 8×10^9 K, while contraction continues until the central core is essentially a ball of neutrons having a radius of about 10 km and a density of 10^{14} g cm^{-3}. At the same time the outer shell of the star is blasted away in an explosion known as a *supernova*.

Since 26Fe56 has the highest binding energy per nucleon of any nuclide, there are no exothermic processes which can lead to the formation of heavier elements. Fusion into heavier species is also precluded by the electrostatic repulsion of the highly charged nuclei. However, the process of neutron capture can still take place (this is the same process that is used to make synthetic elements). The neutrons are by-products of a large variety of stellar processes, and are present in a wide range of energies. Two general types of neutron capture processes are recognized. In an "s" (slow) process, only a single neutron is absorbed and the product usually decomposes by b-decay into a more proton-rich species.

Elements heavier than iron

$$ {}_{26}Fe^{56} + {}_0n^1 \rightarrow {}_{26}Fe^{56} \rightarrow {}_{26}Fe^{59} \rightarrow {}_{27}Co^{59} + {}_{-1}e^0 $$

This process occurs at rates of about 10^5 yr^{-1} , and accounts for the lighter isotopes of many elements. The other process (the "r", or rapid process) occurs in regions of high neutron density and involves multiple captures at rates of 0.1-10 sec^{-1}:

$$ {}_{26}Fe^{56} + 13\ {}_0n^1 \rightarrow {}_{26}Fe^{56} \rightarrow {}_{27}Co^{59} + {}_{-1}e^0 $$

This mechanism favors the heavier, neutron-rich isotopes and the heaviest elements.

Other elements

A few nuclei are not accounted for by any of the processes mentioned. These are all low-abundance species, and they probably result from processes having low rates. Examples are Sn^{112} and Sn^{114}which may be produced through proton-capture, and H^2, Li^6,

Li^7, Be, B^{10} and B^{11}, which may come from spallation processes resulting from collisions of cosmic ray particles with heavier elements

Formation of the solar system

The solar system is believed to have formed about 5 billion years ago as a result of aggregation of cosmic dust and interstellar atoms in a region of space in which the density of such material happened to be greater than average. Over 99.8% of this mass, which consisted mostly of hydrogen, collapsed into a proto-sun; the gravitational energy released in this process raised the temperature sufficiently to initiate the hydrogen fusion reactions discussed above.

The planets

The remaining material probably formed a disk that rotated around the sun. As the temperature dropped to around 2000K, some of the most stable combinations of the elements began to condense out. These substances might have been calcium aluminum silicates, followed by the more volatile iron-nickel system, and then magnesium silicates. The further aggregation of these materials, together with the other constituents of the cooling disk, is now believed to be the origin of the planets. Density estimates indicate that the planets closest to the sun are predominantly rocky in nature, and probably condensed first. The outer planets (Uranus, Neptune and Pluto) appear to consist largely of water ice, methane, and ammonia, with a smaller rocky core.

This page titled 2.1.1.1: Composition and Structure of the Earth is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.1.2: Origin of the Elements

The Earth is composed of 90 chemical elements, of which 81 have at least one stable isotope. Most of these elements have also been detected in stars. Where did these elements come from? The accepted scenario is that the first major element to condense out of the primordial soup was helium, which still comprises about one-quarter of the mass of the known universe.

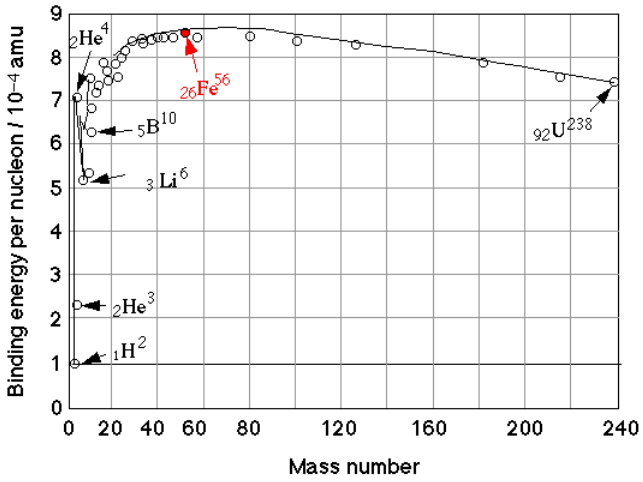


Figure 2.1.1.2.1: Hydrogen is the least thermodynamically stable of the elements, and at very high temperatures will combine with itself in a reaction known as *nuclear fusion* to form the next element, $_{2}He^{4}$. "Heavier" nuclei (that is , those having high atomic numbers, indicated here by the subscript preceding the element symbol), are more stable than "lighter" ones, so this fusion process can continue up to $_{56}Fe$, which is the most energetically stable of all the nuclides. Beyond this point, heavier nuclei slowly become less stable, so fission becomes more likely. FIssion, however, is not considered an important mechanism of primordial *nucleosynthesis*, so other processes are invoked, as discussed farther below.

Primordial Chemistry

According to the *"big bang" theory* for which there is now overwhelming evidence, the universe as we know it (that is, all space, time, and matter) had its origin in a point source or *singularity* that began an explosive expansion about 12-15 billion years ago, and which is still continuing.

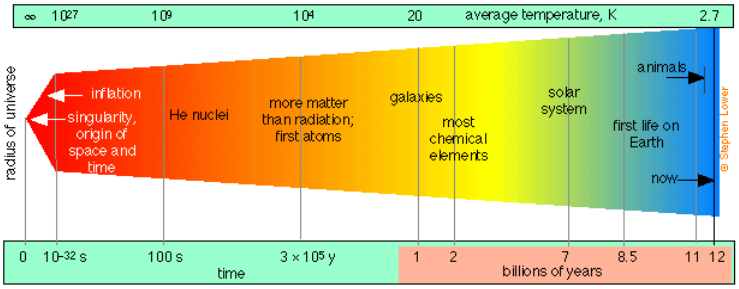


Figure 2.1.1.2.2

Following a brief period of extremely rapid expansion called *inflation*, protons and neutrons condensed out of the initial quantum soup after about 10^{-32} s. Helium and hydrogen became stable during the first few minutes, along with some of the very lightest nuclides up to $_{7}Li$, which were formed through various fusion and neutron-absorption processes. Formation of most heavier elements was delayed for about 10^{6} years until nucleosynthesis commenced in the first stars. Hydrogen still accounts for about 93% of the atoms in the universe.

Stellar Nucleosynthesis

All elements beyond hydrogen were formed in regions where the concentration of matter was large, and the temperature was high; in other words, in stars. The formation of a star begins when the gravitational forces due to a large local concentration of hydrogen bring about a contraction and compression to densities of around 10^{5} g cm^{-3}. This is a highly exothermic process in which the

gravitational potential energy is released as heat, about 1200 kJ per gram, raising the temperature to about 10^7 K. Under these conditions, hydrogen nuclei possess sufficient kinetic energy to overcome their electrostatic repulsion and undergo nuclear fusion:

$$4\,{}^{1}_{1}\mathrm{H} \longrightarrow {}^{4}_{2}\mathrm{He} + 2\,\beta^{+} + 2\,\gamma^{+}\ 2\,\mathrm{n} \qquad (2.1.1.2.1)$$

Hydrogen burning

There will be a net mass loss in above process, which will therefore be highly exothermic and is known as “hydrogen burning”. As hydrogen burning proceeds, the helium collects in the core of the star, raising the density to 10^8 g cm^{-3} and the temperature to 10^8 K. This temperature is high enough to initiate *helium burning*, which proceeds in several steps:

$$2\,{}^{4}_{2}\mathrm{He} \longrightarrow 4\,\mathrm{Be}_{8} + \mathrm{g} \qquad (2.1.1.2.2)$$

The first product, ${}^{8}_{4}\mathrm{Be}$ has a half life of only 10^{-16} sec, but a sufficient amount accumulates to drive the following two reactions:

$${}^{8}_{4}\mathrm{Be} + {}^{4}_{2}\mathrm{He} \longrightarrow {}^{12}_{6}\mathrm{C} + \gamma \qquad (2.1.1.2.3)$$

$${}^{12}_{6}\mathrm{C} + {}^{1}_{1}\mathrm{H} \longrightarrow {}^{13}_{1}\mathrm{N} + {}^{13}_{1}\mathrm{C} + \beta^{+} + \gamma \qquad (2.1.1.2.4)$$

The size of a star depends on the balance between the kinetic energy of its matter and the gravitational attraction of its mass. As the helium burning runs its course, the temperature drops and the star begins to contract. The course of further nucleosynthesis and the subsequent fate of the star itself depends on the star’s mass.

Small stars

If the mass of the star is no greater than 1.4 times the mass of our sun, the star collapses to a white dwarf, and eventually cools to a dark, dense dead star.

Big stars

In larger stars, the gravitational forces are sufficiently strong to overcome the normal repulsion between atoms, and so gravitational collapse continues. The gravitational energy released in this process produces temperatures of 6 $\rightarrow$ 10^8 K, which are sufficient to initiate a complex series of nuclear reactions known as the *carbon-nitrogen cycle*. The net reaction of this cycle is the further fusion of hydrogen to helium, in which C^{12} acts as a catalyst, and various nuclides of nitrogen and oxygen are intermediates. The temperature is sufficiently high, however, to initiate fusion reactions of some of these intermediates:

$_{6}C^{12} + {}_{6}C^{12} \rightarrow {}_{10}Ne^{20} + {}_{2}He^{4}$

$2\ {}_{8}O^{16} \rightarrow {}_{14}Si^{28} + {}_{2}He^{4}$

$2\ {}_{8}O^{16} \rightarrow {}_{16}S^{31} + {}_{0}n^{1}$

Supernovas

The intense gamma radiation that is produced in some of these reactions breaks some of the product nuclei into smaller fragments, which can then fuse into a variety of heavier species, up to the limit of ${}_{26}Fe^{56}$, beyond which fusion is no longer exothermic. The greater relative abundance of elements such as ${}_{6}C^{12}$, ${}_{8}O^{16}$, and ${}_{10}Ne^{20}$ which differ by a ${}_{2}He^{4}$ nucleus, reflects the participation of the latter species in these processes. These exothermic reactions eventually produce temperatures of 8 $\rightarrow$ 10^9 K, while contraction continues until the central core is essentially a ball of neutrons having a radius of about 10 km and a density of 10^{14} g cm^{-3}. At the same time the outer shell of the star is blasted away in an explosion known as a *supernova*.

Note

Only six supernovas have been observed in our galaxy. The supernova of 1987 was the most recent; the one before this occurred in 1604, prior to the invention of the telescope. Tycho Brahe's observation of a supernova in 1572 was crucial in overturning the Aristotelian tradition of the immutability of the "fixed stars", or "firmament". The remains of these supernovas have been detected and studied by X-ray observations.

Thus all of the elements in our solar system that are heavier than iron are the recycled remnants of former stars.

The Chandra X-ray observatory page on **supernovas**

Elements heavier than iron

Since $_{26}Fe^{56}$ has the highest binding energy per nucleon of any nuclide, there are no exothermic processes which can lead to the formation of heavier elements. Fusion into heavier species is also precluded by the electrostatic repulsion of the highly charged nuclei. However, the process of *neutron capture* can still take place (this is the same process that is used to make synthetic elements). The neutrons are by-products of a large variety of stellar processes, and are present in a wide range of energies. Two general types of neutron capture processes are recognized. In an “s” (slow) process, only a single neutron is absorbed and the product usually decomposes by b-decay into a more proton-rich species:

$$_{26}Fe^{56} + {}_{0}n^{1} \rightarrow {}_{26}Fe^{56} \rightarrow {}_{26}Fe^{59} \rightarrow {}_{27}Co^{59} + {}_{-1}e^{0}$$

This process occurs at rates of about 10^{5} yr^{-1} , and accounts for the lighter isotopes of many elements. The other process (the “r”, or rapid process) occurs in regions of high neutron density and involves multiple captures at rates of 0.1-10 sec^{-1}:

$$_{26}Fe^{56} + 13\ {}_{0}n^{1} \rightarrow {}_{26}Fe^{56} \rightarrow {}_{27}Co^{59} + {}_{-1}e^{0}$$

This mechanism favors the heavier, neutron-rich isotopes and the heaviest elements.

Other elements

A few nuclei are not accounted for by any of the processes mentioned. These are all low-abundance species, and they probably result from processes having low rates. Examples are Sn^{112} and Sn^{114} which may be produced through proton-capture, and H^{2}, Li^{6}, Li^{7}, Be, B^{10} and B^{11}, which may come from spallation processes resulting from collisions of cosmic ray particles with heavier elements.

Contributors

- Stephen Lower, Professor Emeritus (Simon Fraser U.) Chem1 Virtual Textbook

This page titled 2.1.1.2: Origin of the Elements is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.1.3: Formation and evolution of the Earth

The solar system is believed to have formed about 5 billion years ago as a result of aggregation of cosmic dust and interstellar atoms in a region of space in which the density of such material happened to be greater than average. Over 99.8% of this mass, which consisted mostly of hydrogen, collapsed into a proto-sun; the gravitational energy released in this process raised the temperature sufficiently to initiate the hydrogen fusion reactions discussed above.

The planets

The remaining material probably formed a disk that rotated around the sun. As the temperature dropped to around 2000K, some of the most stable combinations of the elements began to condense out. These substances might have been calcium aluminum silicates, followed by the more volatile iron-nickel system, and then magnesium silicates. The further aggregation of these materials, together with the other constituents of the cooling disk, is now believed to be the origin of the planets. Density estimates indicate that the planets closest to the sun are predominantly rocky in nature, and probably condensed first. The outer planets (Uranus, Neptune and Pluto) appear to consist largely of water ice, methane, and ammonia, with a smaller rocky core.

Formation of the Earth

The Earth formed by accretion of solid and particulate material that remained after the much more massive amounts of hydrogen and helium present in the original protoplanets had been dispersed out of the solar system. Gradually, the heat produced by decay of radioactive elements brought about partial melting of the silicate rocks; these lower density molten materials migrated upward, leaving the more dense, iron-containing minerals below. This process, which took about 2 million years, was the first of the three stages into which the chemical evolution of the earth is usually divided:

1. **Primary differentiation** of the elements between the core and mantle.
2. **Secondary differentiation** of the elements, reflecting relative ionic sizes, bonding properties, and solubilities (influencing phase behavior such as fractional crystallization, etc.)
3. **Tertiary differentiation**, still operative, involving the interaction of the crust with the hydrosphere and atmosphere.

The above listing should not be taken too literally; all three kinds of processes have probably proceeded simultaneously, and over a number of cycles. Since the earth is losing approximately four times as much heat as is generated by radioactive decay, the principal driving force of primary and secondary differentiation has gradually slowed down. Partial melting of the upper mantle brought about further fractionation as silicon-containing materials of low density migrated outward to form a crust. In its early stages the stronger granitic rocks had not yet appeared, and the crust was mechanically weak. Upwelling flows of lava would break the surface, and the weight of the solidified lava would cause the crust to subside. In some places, magma would solidify underground, forming low-density rock (batholiths) that would eventually rise by buoyancy and push up overlying crust. These mountain-building periods probably occurred in 6-8 major episodes, each lasting about 800 million years.

Rain

At the same time, outgassing of solids released large amounts of HCl, CO, CO_2, H_2S, CH_4 , SO_2 , and SO_3 into the primitive atmosphere. Large amounts of water were present in the primeval rocks in the form of hydrates, which were broken down as the result of the heating. Eventually when the outer crust cooled enough to permit condensation of the water vapor as rain, a new stage of chemical evolution began. The rain was initially highly acidic, equivalent to about 1M HCl; this reacted readily with the basic rocks having high contents of K, Na, Mg, and Ca, leaching them away and forming what would eventually evolve into the oceans. The partial dissolution of the rocks also resulted in large amounts of sediments, which played their own role in the transformation of the earth's surface.

The continents

Within the crust, the lighter materials, being in isostatic equilibrium with the upper mantle, floated higher, and gradually became the nuclei of continents, which grew by accumulating similar material around their boundaries. This picture of continental development is supported by isotopic ratio studies which indicate that the nucleus of the North American continent, the Canadian Shield, is over 2.5 billion years old, while the peripheral parts are less than 0.6 billion years of age.

Primary differentiation of the elements

The more traditional geochemical view of primary differentiation begins with the assumption that the core of the earth is in a chemically reduced state, while the metallic elements constituting the mantle are almost entirely oxidized to their lower free energy cationic forms. Oxygen and sulfur acted as the major electron acceptors in this process, but the abundance of these elements was insufficient to oxidize much of the nickel or iron.

Iron as a reductant

Iron itself is believed to have played a crucial role in the primary differentiation of other metals and of oxidized metallic elements that iron is able to reduce. As the dense molten iron migrated in toward the core, it dissolved (formed a liquid alloy with) any other metals with which it came in contact, and it reduced (donated electrons to) those metallic cations that are less “active” metals than iron under these conditions. The resulting metal would then mix with more of the migrating liquid iron, and be carried along with it into the core.

Redox power of the elements

Accordingly, elements whose reduction potentials are more positive than iron (i.e., are lower-free energy electron sinks) are called *siderophiles*; these elements have a low abundance in the crust and upper mantle. The other two important classes of solid-forming elements are *lithophile* and *chalcophile* (see below.) These generally have more negative reduction potentials than iron, and are distinguished mainly by their relative affinity for oxygen or sulfur. The chalcophiles, of which Cu, Cd, and Sb are examples, tend to form larger, more polarizable ions which can associate with the sulfide ion. The lithophiles comprise those elements such as K, Al, Mn, and Si, which have smaller ions and which combine preferentially with oxygen. This broad classification is reflected in the dominant forms in which many of these elements occur in nature.

Secondary differentiation of the elements

The differential distribution of the elements within one of the main regions of the earth has been studied in detail only in that portion that is accessible, namely the upper crust. It is clear that fractional crystallization from the cooling magma has played an important role. The relative temperatures at which minerals crystallize is determined in large part by their lattice energies, which are in turn related to ionic sizes and charges. Minerals with small, highly charged ions will have higher melting points and should crystallize first. Thus the sodium-containing feldspar albite ($NaAlSi_2O_8$) is found nearer the surface than is its calcium analog anorthite ($CaAlSi_2O_8$). The less abundant elements often do not form minerals of their own, but may replace the ion of a more abundant mineral in its crystal lattice. This is known as *isomorphous replacement*, and it naturally depends on the relative ionic radii. Some ion pairs that undergo isomorphous replacement in minerals are K^+ and Ba^{2+}, Si^{4+} and Ge^{4+}.

Phase behavior

The Phase Rule can be invoked to explain in a very rough way the differentiation of the elements into distinct solid phases.

$P = C + 2 - F$

Taking the degrees of freedom as 2 (fixed temperature and pressure), the six major elemental components (O, Si, Al, Fe, Mg and Na) can form up to six phases. Actually, more than 99% of igneous rocks comprise seven principal mineral phases. These are: the silica minerals, feldspars, feldspathoids, olivine, pyroxenes, amphiboles and micas.

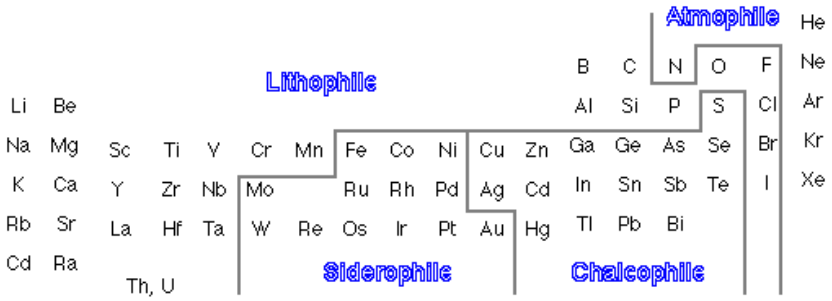
The differential deposition of minerals is also influenced by the temperature-composition phase relations as exemplified by the ordinary two-component phase diagram. If the mineral that is rich in one component and which first crystallizes out is also more dense, then the richer ore will occur near the bottom of the deposit, while a more mixed ore (approaching the eutectic) will remain near the top.

Geochemical classification of the elements

Whether an element is concentrated in the crust or elsewhere depends on its chemical behavior and on the physical properties of its stable compounds. Geochemists have found it convenient to establish the following general classifications:

- *lithophiles* ("rock-loving") elements are those such as Fe, Al, and Si which tend to occur as oxides (and to a lesser extent as chlorides and carbonates.) Elements in this, the largest of all the groups, are concentrated in the crust.

- *chalcophiles* also occur in the crust, but mainly in combination with sulfur and the other chalcogen elements of Groups 15-16.
- *siderophiles* refer to the elements such as Ni which have concentrated in the core along with Fe.
- *atmosphiles* consist of N, H and their volatile compounds
  and the noble gases which concentrate in the atmosphere.



References

"TheTalk" site contains numerous links to discussions of geochronology and other evidence for the age of the earth, and on the pseudoscientific concept of a "young Earth".

Unravelling geological time (Columbia University)

Formation of the Earth's core - description of theories based on high pressure studies

Page last modified: 21.01.2008

© 1998 by Stephen Lower

For information about this Web site or to contact the author,
please see the*Chem1 Virtual Textbook home page*.

SOME RIGHTS RESERVED

This work is licensed under a Creative Commons Attribution-NonCommercial 2.5 License.

This page titled 2.1.1.3: Formation and evolution of the Earth is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.1.4: The Earth's crust

The structure and composition of the outer part of the lithosphere has been profoundly affected by interactions with the atmosphere over one-quarter of the surface area of the earth, and with the hydrosphere over the remaining area. Further modification of the outermost parts of the crust has occurred as the result of the activities of living organisms. These changes have transformed much of the outermost parts of the crust into an unconsolidated surface region called the regolith. Further weathering and translocation of soluble substances often results in a sequence of horizons consisting of sediments, soils, or evaporites.

Chemically, the earth's crust consists of about 80 elements distributed in approximately 2000 compounds or minerals, many of which are of variable composition. Over 99% of the mass of the crustal material is made up of only eight of these elements, however:

Table 2.1.1.4.1: Average amounts of elements in crustal rocks, mg/g.

O	466,000	Na	28,300
Si	277,200	K	25,900
Al	81,300	Mg	20,900
Fe	50,000	Ti	4,400
Ca	36300	H	1,400

The crust has its origin in the upwelling convection currents that bring mantle material near to the surface at the mid-ocean ridges. The reduced pressure causes it to melt into *magma*. The magma may solidify before it reaches the surface, forming basalt, or it may energe from the surface in a volcanic eruption. The oceanic crust consists mostly of the simpler silicate minerals, which are said to be *basic* or *mafic*. The more evolved, silicon-rich rocks found in the continental crust are known as *acidic* or *sialic*.

Oceanic crust is continually being extruded from regions of the plastic mantle that intrude upward to just beneath the ocean's floor at the mid-ocean ridges. A corresponding amount of this crust is being returned to the lithosphere at subduction zones off the West coasts of the Americas, and in the process pushing up the mountain ranges that lie along these coasts. The subducted oceanic crust is reheated and combined with sedimentary material to undergo partial remelting and reworking; this is believed by some to be the origin of *granite*. Subduction proceeds at a rate of a few cm per year, and the complete cycle time is on the order of a few hundred million years.

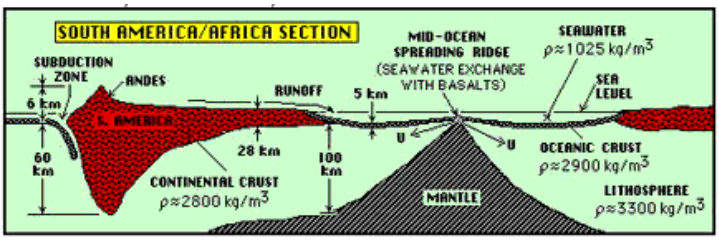


Figure 2.1.1.4.1: The continental and oceanic crusts. This schematic cross section runs in a west-to-east direction, from just of the West coast of South America on the left to the African continent on the right. Oceanic crust is shown in black, with with continental crust in red. (From Cosmos and Earth Study Guide, NYU)

Both oceanic and contintental crusts float on the more dense upper lithosphere, and gradually shift their positions as they push against each other, and in response to the slow convective motions in the medium that supports them. The **continental crust** is thicker than the oceanic crust, but it is also less dense, which allows it to float higher (and thus to differentiate continents from oceans.) The lower density also prevents it from being subducted. Recycling can occur indirectly as continental material erodes and

is deposited as sediments on the ocean floor, but this is a much slower process and one that takes billions instead of millions of years. Some of the very oldest rocks, found in Greenland and Labrador, have been dated at 3.9 billion years, and thus approach the age of the Earth itself.

Chemistry of the crust

When magma crystallizes it forms igneous rock, the major component of the Earth's crust. The crystallization is a complex process which is not entirely understood, due largely to the lack of sufficient thermodynamic data on the various components at high temperatures and pressures. It is known that the different components of magma have differing melting points and densities, and that the phase behavior of multicomponent systems based on some of these substances is quite complex, involving binary and ternary eutectics, solid solutions, the presence of dissolved water (under pressure), and incongruent melting. One consequence of this complexity is that the composition of the magma will change as crystallization takes place; different substances will crystallize at various stages, and the resulting solids may migrate toward the top or bottom of the region if their densities differ greatly from that of the magma

It is well known that larger crystals form when a melt cools more slowly. This principle affords a simple distinction between the coarser-grained plutonic rocks, which are believed to have been formed by gradual cooling of magma pockets within the crust, and the fine-grained volcanic rocks such as basalt. Under the influence of heat and pressure, particularly at plate boundaries, solid crustal material may undergo partial or complete remelting, followed by cooling and transformation into metamorphic rocks such as gneiss, micas, quartzite, and possibly granites.

> ***Granite*** *was once thought to be an igneous rock, originating from the crystallization of a particular kind of magma. The association of granitic rocks with mountainous regions, and the similarity of their compositions in widely scattered regions, lends credence to the more recent hypothesis that granitic rocks are of metamorphic origin.*

Another class of rock is *sedimentary rock*, formed from the consolidation of material produced by weathering and other chemical, and biological processes. Sedimentary rocks cover about three-quarters of the land area of the earth; 80% are shales, 15% sandstones and 5% limestones.

Composition of rock

The chemical composition of rocks tends to be complex and variable, and can only be specified in a precise way at the structural level. The traditional way of expressing rock compositions is in terms of the mass percent of the oxides of the elements present in the rock.

Table 2.1.1.4.1: **: Chemical composition of a typical rock** (quartz-feldspar-biotite gneiss). The figures are in percent by weight.

oxide	common name	fresh rock	weathered rock
SiO_2	silica	71.54	70.30
Al_2O_3	alumina	14.62	18.34
Fe_2O_3	ferric oxide	0.69	1.55
FeO	ferrous oxide	1.64	0.22
MgO	magnesia	0.77	0.21
CaO	lime	2.08	0.10
Na_2O	soda	3.84	0.09
K_2O	potash	0.32	5.88
H_2O	water	0.32	5.88
others		0.65	0.54
total		100.07	99.70

oxide	common name	fresh rock	weathered rock
total			

This does not mean that these oxides, or the structural units they represent, are actually present as such in a rock. In the chemical analysis of rocks, oxygen is generally not determined separately. When it is, however, it is found in an amount that would be expected to combine stoichiometrically with the other elements present. Thus the composition of albite can be written as either $NaAlSi_3O_8$ or $Na_2O{\cdot}Al_2O_3{\cdot}6SiO_2$. Some rocks contain varying ratios of certain elements. For example olivine, which can be considered a solid solution of Mg_2SiO_4 and Fe_2SiO_4, can be represented by $(Mg,Fe)_2SiO_4$; this implies that the ratio of metal to silica is constant, and that magnesium is ordinarily present in greater amount than iron.

The major structural elements of rock (both in the crust and in the mantle) are the silicate minerals, built from silicon atoms surrounded tetrahedrally by four oxygens. The simplest of these consist just of SiO_4^{4-} tetrahedra interspersed with positive ions to achieve electroneutrality; olivine, $(Mg,Fe)_2SiO_4$ is a well known example. More commonly, the silicate groups polymerize by sharing one or more oxygen atoms at adjacent tetrahedral corners. Depending on the number of joined corners per silicate unit, this can lead to the formation of a wide variety of chains (pyroxenes, amphiboles) and sheets (micas), culminating in the complete tetrahedral polymerization that produces quartz, SiO_2.

Higher degrees of polymerization are associated with higher ratios of Si to O, smaller quantities of positive ions, and higher melting points. Thus when magma cools, the first silicates to crystallize are the olivines, followed by chain and sheet minerals having progressively higher degrees of polymerization and smaller fractions of cations of metals such as Fe and Mg.

Distribution of elements; ores

Although some elements are distributed fairly uniformly throughout the crust, others occur at greatly enhanced concentrations in localized areas. There are two general processes that result in these localized excesses, which are called *ores* when their extraction and refining is economically feasible

The first of these relates to how well a metallic ion can fit into the silicate lattice structure. Ions having the right charge and size can readily enter this structure, displacing the more common ions of Fe, Al and Mg. Such ions (of which Ga^{3+} is an example) are readily soluble in other minerals and thus are widely disributed and do not concentrate into ores. Other ions may be too large (Cs and La), too small (Li, Be, B) or too highly charged (Nb, Ta, W) to be accommodated in silicate mineral structures; these elements tend to remain in the magma as it solidifies, finally forming solid minerals only in the last stages of cooling.

The other major source of ores is *hydrothermal* formation. Magma contains some water itself, and additional water from the surface is able to reach the heated rock near magma chambers. At the very high temperatures and pressures that prevail in these regions, the water can dissolve many compounds such as sulfides which are normally considered highly insoluble. When these superheated solutions rise to the surface the solids are re-deposited, often in highly concentrated form. Ores of Cu, Sn, W, and possibly some iron ores, as well as some native metals such as gold, are believed to be formed in this way.



Figure 2.1.1.4.2: The hydrothermal vent photographed here emits both black and white "smoke".

Hydrothermal vents known as “black smokers” have been observed at sites of sea-floor spreading; the “smoke” consists of metallic sulfides which precipitate in the cold seawater. The veins of pyrites (FeS_2) and similar sulfide minerals that are often observed in rock formations are the result of hydrothermal solutions that once penetrated cracks and fissures in the rock.

Deep-sea hydrothermal vents (many more pictures) Theories of ore formation (a more technical tutorial)

Chemical weathering

The weathering of rocks at the earth’s surface is a complex process involving both physical and chemical changes. The latter tend in principle to be rather simple kinds of reactions involving dissolution, reaction with carbon dioxide, hydrolysis, hydration, and oxidation. The difficulty in studying them and in arriving at a quantitative description is that these reactions occur very slowly and may never reach an equilibrium state. A comparison of the two rightmost columns in Table 2 on page 14 provides some illustration of the overall effect of these changes, although it must be emphasized that these are relative composition data, and thus cannot show how much of a given component has been lost. In general, sodium, calcium and magnesium seem to be lost more rapidly than potassium and silicon, while iron and aluminum decrease very slowly. Individual rates are of course dependent on the particular structural units containing the element, and also vary somewhat with grain size and condition of the surface.

Action of water

Water is undoubtedly the most important weathering agent. Not only does it act as a solvent for ionic dissolution products, but it also brings other active agents such as carbon dioxide and oxygen into intimate contact with the rock material. As water percolates into the outermost layers of the crust, it extends the zone of weathering beneath the surface; the effects of this are quite noticeable in a number of buried sedimentary materials such as Paleozoic sandstones, which tend to be depleted of all but the most resistant minerals. Dissolution, the simplest of all the weathering processes, usually results in ionic species, some of which may react with water to yield acidic or alkaline solutions. Dissolution of silica, however, results in the neutral species H_4SiO_4. Reactions involving hydration and dehydration are very common, and since the free energy changes tend to be small, these reactions can usually take place in either direction under slightly different conditions. Thus gypsum and anhydrite are interconvertable at observable rates under common environmental conditions:

$CaSO_4 \cdot 2H_2O \rightarrow CaSO_4 + 2\ H_2O$

In many cases, however, the reaction products are not very well characterized, thermodynamic data is lacking, and the reactions proceed so slowly that they are not entirely understood. For example, both hydrous and anhydrous iron oxides can be found in similar geologic environments, but the little is known about the interconversion process, represented approximately as

$Fe_2O_3 + H_2O \rightarrow 2\ FeOOH$

Solid carbonates tend to dissolve in acidic solutions, including those produced when atmospheric carbon dioxide dissolves in water. Thus the major surface limestone deposits (largely $CaCO_3$, with some admixture of $MgCO_3$) tend to be highly eroded in non-arid

regions, and the local groundwater may have Ca^{2+} as high as 0.1-0.2M. Thermodynamics can unambiguously predict the most stable oxidation state of a metal ion under given conditions of pH and oxidant concentration. The mechanisms tend to be very uncertain, however. For one thing, both the reactant and product can often exist in various states of hydration, and the dissolved species (which probably undergo the actual oxidation) often consist of polycations and complexed species.

Oxidation of iron

Compounds of Fe(II), for example, will always tend to oxidize to Fe(III) in the presence of air; the various oxides of iron are responsible for the bright colors seen in many geological formations, and in certain soils. Some of the net reactions that probably occur are

$Fe_2SiO_4 + 1/2\ O_2 + H_2O \rightarrow Fe_2O_3 + H_2SiO_4$

$2\ FeCO_3 + 1/2\ O_2 \rightarrow Fe_2O_3 + 2\ H_2CO_3$

An environmental side effect of the first process is the release of hydrated silica. Also, where both starting materials are present, the carbonic acid produced in the second reaction is believed to promote the dissolution of ferrous silicate, creating a source of Fe(II) ions that can be rapidly oxidized:

$Fe_2SiO_4 + 4\ H_2CO_3 \rightarrow 2\ Fe^{2+} + 4\ HCO_3^- + H_4SiO_4$

$Fe^{2+} + 4\ HCO_3^- + 1/2\ O_2 \rightarrow Fe_2O_3 + 4\ H_2CO_3$

The oxidation of sulfides can produce strongly acidic solutions:

$2\ FeS_2 + 15/2\ O_2 + 4\ H_2O \rightarrow Fe_2O_3 + 2\ SO_4^{2-} + 8\ H^+$

The effects of this can be seen in formations containing outcrops of pyrite veins, where the surrounding rocks are heavily stained with yellow and brown Fe(III) oxides, and the groundwater tends to be highly acidic. This process is mediated by microorganisms, and is an important source of acid pollution associated with mines and mine tailings.

Sequence of weathering

The various components of rocks weather at different rates. The more basic components such as CaO and MgO tend to disappear first, especially if in contact with groundwaters containing high CO2 concentrations. For rocks in general, the first reaction is usually hydration, followed by hydrolysis which can be summarized by

$4\ KAlSi_3O_8(s) + 22\ H_2O \rightarrow Al_4Si_4O_{10}(OH)_8(s) + 8\ H_4SiO_4(aq) + 4\ K^+ + 4\ OH^-$

in which other Group 1 or 2 cations might replace potassium. The product $Al_4Si_4O_{10}(OH)_8$ is *kaolinite*, a form of clay (see below).

In general, the rocks which crystallized first from the magma (the Ca-feldspars and olivines) weather more rapidly than do the lower-melting rocks.

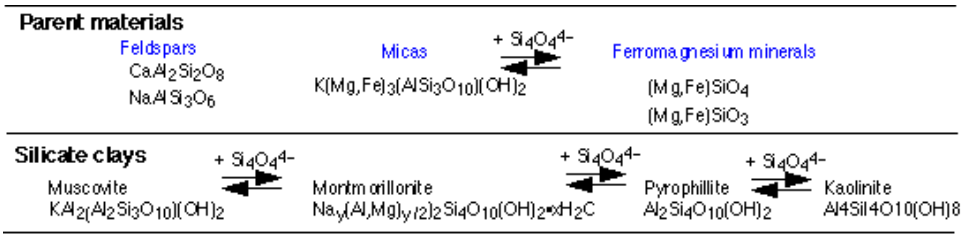


Figure 2.1.1.4.3:

Clays

Clays are the solid end products of the weathering of rocks. They are basically composed of alternating sheets of "SiO_2" and "AlO_6" units in ratios of 1:1 (kaolinite), 2:1 (montmorillonite and vermiculite) and 2:2 (chlorite). In between the sheets, and

holding them together by hydrogen bonding are water molecules. Also present are cations such as K^+, Ca^{2+} and Mg^{2+} which act to neutralize the negative charges of the oxide ions.

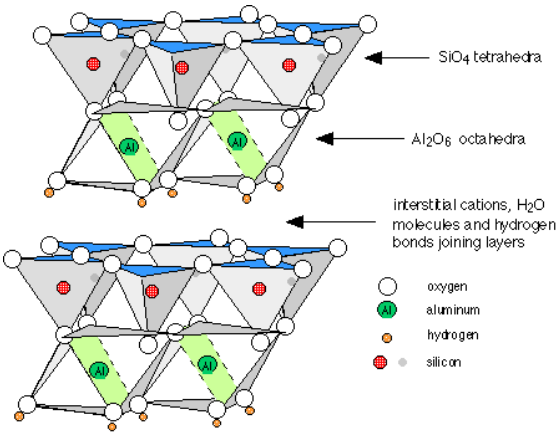


Figure 2.1.1.4.4: Structure of a clay, showing two layers of the stacked sheets of kaolinite

Physical weathering

The major agents of physical weathering of exposed rocks are rapid changes in temperature (promoting fracture by differential expansion), the abrasive action of windborne material and glacier movement, and especially by the penetration of water into cracks and its subsequent freezing. The expansion of water on freezing can exert a pressure of 150 kg cm^{-2}, whereas the tensile strength of a typical rock is around 120 kg cm^{-2} . The roots of some plants are able to penetrate rock quite effectively, producing comparable expansive pressures in subsurface rocks.

Composition and structure of soils

Soils are a product of the interaction of water, air, and living organisms with exposed rocks or sediments at the earth's surface. A typical soil contains about 45% inorganic solids and 5% organic solids by volume. Water and air each make up about 20-30%.

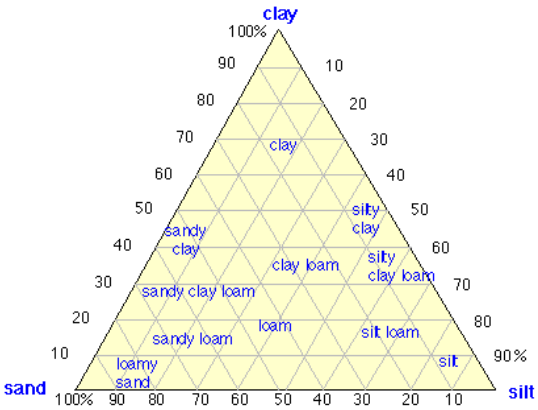


Figure 2.1.1.4.5: A simple way of classifying soils is based on the relative quantities of clay, silt and sand in the solid component. For ordinary agricultural purposes, loams are considered the most ideal soil type.

Mineral Components of soil

The primary inorganic components of soils consist of sand and silt particles that come directly from the parent rocks. This fraction is dominated by quartz and feldspars (aluminosilicates). Secondary components are formed by chemical changes within the soil itself, or in sediments from which the soil derived. These are most commonly clays, but may also include calcite, gypsum, and sulfide minerals such as pyrites; the latter are formed by bacterial action under reducing conditions in the presence of organic matter.

The clays have an especially important effect on both the physical properties of the soil, and on its ability to store plant nutrients, including trace nutrients such as Mo and Mn. These properties are due to the high ion-exchange capacity of clays. The more highly charged cations such as Al3+ and Fe3+ tend to be more strongly absorbed within the inter-sheet regions than do Mg^{2+} or K^{+}. As plants withdraw these latter cations from the soil water, more are released by the clay components, which thus act as nutrient reservoirs.

The ion-exchange properties of clays also help to maintain the pH balance of soils, through the exchange of H^{+} and cations such as Ca^{2+}. The soil pH, in turn, strongly affects the solubility of nutrient cations, and thus their availability to plants. For example, the uptake of phosphorus (in the form of $H_2PO_4^{2-}$, is only efficient within the rather narrow pH range between 6 and 7. Below 6, dihydrogen phosphates of Fe and Al are precipitated, while insoluble $Ca_3(PO_4)_2$ forms at higher pH's.

Organic Components

Part of the organic matter of soil consists of organisms (mainly bacteria and fungi) and roots and root hairs. The remainder is largely in the form of fulvic and *humic* acids. These substances of indefinite composition are classified on the basis of their solubility behavior; fulvic acids remain in solution at pH 2, but humic acids, having molecular weights of 20,000 to 100,000, are precipitated. Both are flexible polyelectrolytes that interact strongly with their own kind and with inorganic ions.

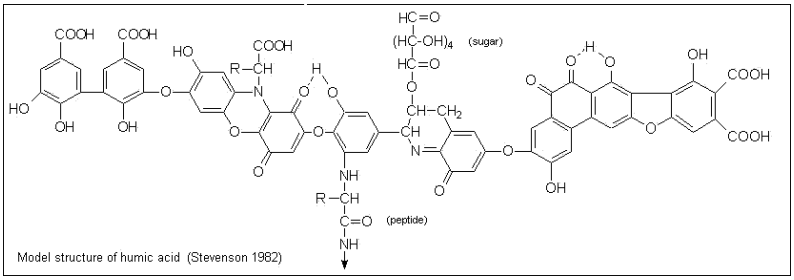


Figure 2.1.1.4.6: Model structure of humic acid (Stevenson 1982)

Associated with the fulvic and humic fractions are a wide variety smaller molecules such as alkanes, amino acids, amino sugars, sulfur and phosphorus derivatives of sugars, etc. Part of the organic carbon in a fertile soil is recycled in 1-2 years; plant residues, which are the major source of soil organic matter, have a half-life of days to months. Once carbon gets incorporated into humic substances, it is locked into a much slower recycling process; the turnover times of fulvic acids are a hundred years or more, while those for fulvic acids are around a thousand years. For this reason, humic substances are the major reservoir of organic carbon in soil. Organic matter, particularly polysaccharides, binds strongly to the cation components of clay colloids; the two together act as cementing agents and strongly influence the consistency and structure of the soil.

Water

Soil water is held by capillary action and adsorption with varying degrees of tenacity. This water binding strength is traditionally expressed in terms of the pressure, or "tension" that would be required to force the water out of the soil. The tension of capillary water varies over a wide range of 0.1-32 atm; only in the lower half of this range will it be available to plants, which can exert an osmotic pressure of up to about 15 atm. Water in excess of the capillary capacity fills larger voids and is called gravitational water. Its presence in surface soils corresponds to a flooded condition that inhibits plant growth by reducing soil aeration.

Air

The gas phase within soil pores generally has a CO_2 content of 5-50 times that of the atmosphere due to the action of organisms. O_2 tends to be depleted to roughly the extent that CO_2 is present in excess. Under conditions of poor aeration (i.e., exchange with the atmosphere), considerable quantities of N_2O, NO, H_2, CH_4, C_2H_4, and H_2S may also be present.

What is soil? An introductory survey

Soils interactive textbook

Page last modified: 21.01.2008

© 1998 by Stephen Lower

For information about this Web site or to contact the author,
please see the*Chem1 Virtual Textbook home page*.

This work is licensed under a Creative Commons Attribution-NonCommercial 2.5 License.

O	466,000
Si	277,200
Al	81,300
Fe	50,000
Ca	36300

This page titled 2.1.1.4: The Earth's crust is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

SECTION OVERVIEW

2.1.2: The Hydrosphere

Water is the most abundant substance at the earth's surface. Almost all of it is in the oceans, which cover 70% of the surface area of the earth. However, the amounts of water present in the atmosphere and on land (as surface runoff, lakes and streams) is great enough to make it a significant agent in transporting substances between the lithosphere and the oceans. Water interacts with both the atmosphere and the lithosphere, acquiring solutes from each, and thus provides the major chemical link between these two realms. The various transformations undergone by water through the different stages of the hydrologic cycle act to transport both dissolved and particulate substances between different geographic locations.

2.1.2.1: Water, Water Everywhere...

2.1.2.2: The hydrosphere and the oceans

2.1.2.3: Chemistry and geochemistry of the oceans

2.1.2.4: Chemical budgets of oceanic elements

This page titled 2.1.2: The Hydrosphere is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.2.1: Water, Water Everywhere...

Water is the most abundant substance at the earth's surface. Almost all of it is in the oceans, which cover 70% of the surface area of the earth. However, the amounts of water present in the atmosphere and on land (as surface runoff, lakes and streams) is great enough to make it a significant agent in transporting substances between the lithosphere and the oceans. Water interacts with both the atmosphere and the lithosphere, acquiring solutes from each, and thus provides the major chemical link between these two realms. The various transformations undergone by water through the different stages of the hydrologic cycle act to transport both dissolved and particulate substances between different geographic locations.

Reservoir	Volume / 10^6, km^3	Percent of total
Oceans	1370	97.25
Ice caps and glaciers	39	2.05
Deep groundwater	5.3	0.38
Shallow groundwater	4.3	0.30
Lakes	0.125	0.01
Soil moisture	0.065	0.005
Atmosphere	0.013	0.001
Rivers	0.0017	0.0001
Biosphere	0.0006	0.00004
Total	**1408.7**	**100**

Composition of Seawater

The composition of the ocean has attracted the attention of some of the more famous names in science, including Robert Boyle, Antoine Lavoisier and Edmund Halley. Their early investigations tended to be difficult to reproduce, owing to the different conditions under which they crystallized the various salts. As many as 54 salts, double salts and hydrated salts can be obtained by evaporating seawater to dryness. At least 73 elements are now known to be present in seawater.

Table 1: Major ions of seawater: These values, expressed in parts per thousand, are for seawater of 35% salinity.

cations	g/kg	anions	g/kg
Na^+	10.77	Cl^-	19.354
Mg^{2+}	1.29	SO_4^{2-}	2.712
Ca^{2+}	0.412	Br^-	0.087
K^+	0.399		
Sr^{2+}	0.0079		
Al^{3+}	0.005		

The best way of characterizing seawater is in terms of its ionic content, shown above. The remarkable thing about seawater is the constancy of its relative ionic composition. The overall salt content, known as the salinity (grams of salts contained in 1 kg of seawater), varies slightly within the range of 32-37.5%, corresponding to a solution of about 0.7% salt content. The ratios of the concentrations of the different ions, however, are quite constant, so that a measurement of Cl^- concentration is sufficient to determine the overall composition and total salinity. Although most elements are found in seawater only at trace levels, marine organisms may selectively absorb them and make them more detectable. Iodine, for example, was discovered in marine algae

(seaweeds) 14 years before it was found in seawater. Other elements that were not detected in seawater until after they were found in marine organisms include barium, cobalt, copper, lead, nickel, silver and zinc. Si^{32}, presumably deriving from cosmic ray bombardment of Ar, has been discovered in marine sponges.

pH balance. Reflecting this constant ionic composition is the pH, which is usually maintained in the narrow range of 7.8-8.2, compared with 1.5 to 11 for fresh water. The major buffering action derives from the carbonate system, although ion exchange between Na+ in the water and H^+ in clay sediments has recently been recognized to be a significant factor.

Conservative and non-conservative substances

The major ionic constituents whose concentrations can be determined from the salinity are known as conservative substances. Their constant relative concentrations are due to the large amounts of these species in the oceans in comparison to their small inputs from river flow. This is another way of saying that their residence times are very large.

Table 2: Replacement time with respect to river addition for some components of seawater. Concentrations in micromols/L

component	concentration in river water	concentration in seawater	residence time, 10^6 y
Cl^-	250	558,000	87
Na^+	315	479,000	55
Mg^{2+}	150	54,300	13
SO_4^{2-}	120	28,900	8.7
Ca^{2+}	367	10,500	1
K^+	36	10,400	10
HCO_3^-	870	2000	0.083
H_4SiO_4	170	100	0.021
NO_3^-	10	20	0.072
$H_2PO_4^-$	0.7	1	0.080

A number of other species, mostly connected with biological activity, are subject to wide variations in concentration. These include the nutrients NO_3^-, NO_2^-, NH_4^+, and HPO_4^{2-}, which may become depleted near the surface in regions of warmth and light. As was explained in the preceding subsection on coastal upwelling, offshore prevailing winds tend to drive Western coastal surface waters out to sea, causing deeper and more nutrient-rich water to be drawn to the surface. This upwelled water can support a large population of phytoplankton and thus of zooplankton and fish. The best-known example of this is the anchovy fishery off the coast of Peru, but the phenomenon occurs to some extent on the West coasts of most continents, including our own.

Other non-conservative components include Ca^{2+} and SiO_4^{2-}. These ions are incorporated into the solid parts of marine organisms, which sink to greater depths after the organisms die. The silica gradually dissolves, since the water is everywhere undersaturated in this substance. Calcium carbonate dissolves at intermediate depths, but may reprecipitate in deep waters owing to the higher pressure. Thus the concentrations of Ca and of SiO_4^{2-} tend to vary with depth. The gases O_2 and CO_2, being intimately involved with biological activity, are also non-conservative, as are N_2O and CO.

Organic matter

Most of the organic carbon in seawater is present as dissolved material, with only about 1-2% in particulates. The total organic carbon content ranges between 0.5 mg/L in deep water to 1.5 mg/L near the surface. There is still considerable disagreement about the composition of the dissolved organic matter; much of it appears to be of high molecular weight, and may be polymeric. Substances qualitatively similar to the humic acids found in soils can be isolated. The greenish color that is often associated with coastal waters is due to a mixture of fluorescent, high molecular weight substances of undetermined composition known as *"Gelbstoffe"*. It is likely that the significance of the organic fraction of seawater may be much greater than its low abundance

would suggest. For one thing, many of these substances are lipid-like and tend to adsorb onto surfaces. It has been shown that any particle entering the ocean is quickly coated with an organic surface film that may influence the rate and extent of its dissolution or decomposition. Certain inorganic ions may be strongly complexed by humic-like substances. The surface of the ocean is mostly covered with an organic film, only a few molecular layers thick. This is believed to consist of hydrocarbons, lipids, and the like, but glycoproteins and proteoglycans have been reported. If this film is carefully removed from a container of seawater, it will quickly be reconstituted. How significant this film is in its effects on gas exchange with the atmosphere is not known.

Regulation of ocean composition

The salinity of the ocean appears to have been about the same for at least the last 200 million years. There have been changes in the relative amounts of some species, however; the ratio of Na/K has increased from about 1:1 in ancient ocean sediments to its present value of 28:1. Incorporation of calcium into sediments by the action of marine organisms has depleted the Ca/Mg ratio from 1:1 to 1:3.

Table: Mass balance of P, C, and Ca for the oceans

element	input to ocean	dissolved in seawater	in dead organisms	loss to sediments	residence time, y
phosphorus	1 (erosion)	1	1	1 (organic)	10,000
carbon	100 as CO_2 500 as carbonate	1000	125	100 (organic) 600 (carbonate)	165,000
calcium	500 (erosion)	5000	25	500 (as $CaCO_3$)	10^6

If the composition of the ocean has remained relatively unchanged with time, the continual addition of new mineral substances by the rivers and other sources must be exactly balanced by their removal as sediment, possibly passing through one or more biological systems in the process.

Where does the Salt Come From?

In 1715 Edmund Halley suggested that the age of the ocean (and thus presumably of the world) might be estimated from the rate of salt transport by rivers. When this measurement was actually carried out in 1899, it gave an age of only 90 million years.

> *This is somewhat better than the calculation made in 1654 by James Ussher, the Anglican Archbishop of Armagh, Ireland, based on his interpretation of the Biblical book of Genesis, that the world was created at 9 A.M. on October 23, 4004 BC, but it is still far too recent, being about when the dinosaurs became extinct. What Halley actually described was the residence time, which is about right for Na but much to long for some of the minor elements of seawater.*

The commonly stated view that the salt content of the oceans derives from surface runoff that contains the products of weathering and soil leaching is not consistent with the known compositions of the major river waters (See Table). The halide ions are particularly over-represented in seawater, compared to fresh water. These were once referred to as “excess volatiles”, and were attributed to volcanic emissions. With the discovery of plate tectonics, it became apparent that the locations of seafloor spreading at which fresh basalt flows up into the ocean from the mantle are also sources of mineral-laden water. Some of this may be seawater that has cycled through a hot porous region and has been able to dissolve some of the mineral material owing to the high temperature. Much of the water, however, is “juvenile” water that was previously incorporated into the mantle material and has never before been in the liquid phase. The substances introduced by this means (and by volcanic activity) are just the elements that are “missing” from river waters. Estimates of what fraction of the total volume of the oceans is due to juvenile water (most of it added in the early stages of mantle differentiation that began a billion years ago) range from 30 to 90%.

Geochemical processes involving the oceans

The oceans can be regarded as a product of a giant acid-base titration in which the carbonic acid present in rain reacts with the basic materials of the lithosphere. The juvenile water introduced at locations of ocean-floor spreading is also acidic, and is partly neutralized by the basic components of the basalt with which it reacts. Surface rocks mostly contain aluminum, silicon and oxygen combined with alkali and alkaline-earth metals, mainly potassium, sodium and calcium. The CO_2 and volcanic gases in rainwater react with this material to form a solution of the metal ion and HCO_3^-, in which is suspended some hydrated SiO^2. The solid material left behind is a clay such as kaolinite, $Al_2Si_2O_5(OH)_4$. This first forms as a friable coating on the surface of the weathered rock; later it becomes a soil material, then an alluvial deposit, and finally it may reach the sea as a suspended sediment. Here it may undergo a number of poorly-understood transformations to other clay sediments such as illites. Sea floor spreading eventually transports these sediments to a subduction region under a continental block, where the high temperatures and pressures permit reactions that transform it into hard rock such as granite, thus completing the geochemical cycle.

Deep-sea hydrothermal vents are now recognized to be another significant route for both the addition and removal of ionic substances from seawater.

Distribution and cycling of elements in the oceans

Although the relative concentrations of most of the elements in seawater are constant throughout the oceans, there are certain elements that tend to have highly uneven distributions vertically, and to a lesser extent horizontally. Neglecting the highly localized effects of undersea springs and volcanic vents, these variations are direct results of the removal of these elements from seawater by organisms; if the sea were sterile, its chemical composition would be almost uniform.

Plant life can exist only in the upper part of the ocean where there is sufficient light available to drive photosynthesis. These plants, together with the animals that consume them, extract nutrients from the water, reducing the concentrations of certain elements in the upper part of the sea. When these organisms die, they fall toward the lower depths of the ocean as particulate material. On the way down, some of the softer particles, deriving from tissue, may be consumed by other animals and recycled. Eventually, however, the nutrient elements that were incorporated into organisms in the upper part of the ocean will end up in the colder, dark, and essentially lifeless lower part.

Mixing between the upper and lower reservoirs of the ocean is quite slow, owing to the higher density of the colder water; the average residence time of a water molecule in the lower reservoir is about 1600 years. Since the volume of the upper reservoir is only about 1/20 of that of the lower, a water molecule stays in the upper reservoir for only about 80 years.

Except for dissolved oxygen, all elements required by living organisms are depleted in the upper part of the ocean with respect to the lower part. In the case of the major nutrients P, N and Si, the degree of depletion is sufficiently complete (around 95%) to limit the growth of organisms at the surface. These three elements are said to be biolimiting. A few other biointermediate elements show partial depletion in surface waters: Ca (1%), C (15%), Ba (75%).

The organic component of plants and animals has the average composition $C_{80}N_{15}P$. It is remarkable that the ratio of N:P in seawater (both surface and deep) is also 15:1; this raises the interesting question of to what extent the ocean and life have co-evolved.

In the deep part of the ocean the elemental ratio corresponds to $C_{800}N_{15}P$, but of course with much larger absolute amounts of these elements. Eventually some of this deeper water returns to the surface where the N and P are quickly taken up by plants. But since plants can only utilize 80 out of every 800 carbon atoms, 90 percent of the carbon will remain in dissolved form, mostly as HCO_3^-.

> *To work out the balance of Ca and Si used in the hard parts of organisms, we add these elements to the average composition of the lower reservior to get $Ca_{3200}Si_{50}C_{800}N_{15}P$. Particulate carbon falls into the deep ocean in the ratio of about two atoms in organic tissue to one atom in the form of calcite. This makes the overall composition of detrital material something like $C_{120}N_{15}P$; i.e., 80 organic C's and 40 in $CaCO_3$. Accompanying these 40 calcite units will be 40 Ca atoms, but this represents a minor depletion of the 3200 Ca atoms that eventually return to the surface, so this element is only slightly*

> *depleted in the upper waters. Silicon, being far less abundant, is depleted to a much greater extent.*

The oceanic sediments

The particulate shower

A continual rain of particulate material from dead organisms falls through the ocean. This shower is comprised of three major kinds of material: calcite ($CaCO_3$), silica (SiO_2), and organic matter. The first two come from the hard parts of both plants and animals (mainly microscopic animals such as foraminifera and radiolarians). The organic matter is derived mainly from the soft tissues of organisms, and from animal fecal material. Some of this solid material dissolves before it reaches the ocean floor, but not usually before it enters the deep ocean where it will remain for about 1600 years.

The remainder of this material settles onto the floor of the sea, where it forms one component of a layer of sediments that provide important information about the evolution of the sea and of the earth. Over a short time scale of months to years, these sediments are in quasi-equilibrium with the seawater. On a scale of millions of years, the sediments are merely way-stations in the geochemical cycling of material between the earth's surface and its interior.

The oceanic sediments have three main origins:

- **Detrital** material is derived largely from particles deposited in the ocean by rivers and also directly by the wind. These materials are mostly aluminosilicates (clays), along with some quartz. These substances accumulate on the deep ocean floor at a rate of about 0.3 g cm^{-2} per 1000 years.
- **Authigenic** materials are formed by precipitation within the ocean or by crystallization within the sediment itself. These constitute only a small fraction of the total sediment.
- **Biogenic** components consist mainly of the calcium carbonate and silica that make up the hard parts of organisms. Scavenging by other organisms is so efficient that only about 0.3 % of the total deep sea sediment consists of actual organic material.

Our main interest lies with the silica and calcium carbonate, since these substances form a crucial part of the biogeological cycle. Also, their distributions in the ocean are not uniform- a fact that must tell us something.

The skeletons of diatoms and radiolarians are the principal sources of silica sediments. Since the ocean is everywhere undersaturated with respect to silica, only the most resistant parts of these skeletons reach the bottom of the deep ocean and get incorporated into sediments. Silica sediments are less common in the Atlantic ocean, owing to the lower content of dissolved silica.

The parts of the ocean where these sediments are increasing most rapidly correspond to regions of upwelling, where deep water that is rich in dissolved silica rises to the surface where the silica is rapidly fixed by organisms. Where upwelling is absent, the growth of the organisms is limited, and little silica is precipitated. Since deep waters tend to flow from the Atlantic into the Pacific ocean where most of the upwelling occurs, Atlantic waters are depleted in silica, and silica sediments are not commonly found in this ocean.

For calcium carbonate, the situation is quite different. In the first place, surface waters are everywhere supersaturated with respect to both calcite and aragonite, the two common crystal forms of $CaCO_3$. Secondly, Ca^{2+} and HCO_3^- are never limiting factors in the growth of the coccoliths (plants) and forams (animals) that precipitate $CaCO_3$; their production depends on the availability of phosphate and nitrogen. Because these elements are efficiently recycled before they fall into the deep ocean, their supply does not depend on upwelling, and so the production of solid is more uniformly distributed over the world's oceans.

More importantly, however, the chances that a piece of carbonate skeleton will end up as sediment will be highly dependent on both the local CO_3^{2-} concentration and the depth of the ocean floor. These factors give rise to small-scale variations in the production of carbonate sediments that can be quite wide-ranging.

Ocean sediments and continental drift

New crust is being generated and moving away from the crests of the mid-ocean ridges at a rate of a few centimetres per year. Although the crests of these ridges are relatively high points, projecting to within about 3000 m of the surface, the continual injection of new material prevents sediments from accumulating in these areas. Farther from the crests, carbonate sediments do

build up, eventually reaching a depth of about 500 m, but by this time the elevation has dropped off below the saturation horizon, so from this point on the carbonate sediments are overlaid by red clay.

If we drill a hole down through a part of the ocean floor that is presently below the saturation horizon, the top part of the drill core will consist of clay, followed by $CaCO_3$ at greater depths. The core may also contain regions in which silica predominates. Since silica production is very high in equatorial regions, the appearance of such a layer suggests that this particular region of the oceanic crust has moved across the equator.

Contributors and Attributions

- Stephen Lower, Professor Emeritus (Simon Fraser U.) Chem1 Virtual Textbook

This page titled 2.1.2.1: Water, Water Everywhere... is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.2.2: The hydrosphere and the oceans

> *How inappropriate to call this planet Earth,*
> *when clearly it is Ocean."*
>
> *Arthur C. Clarke*

Water is the most abundant substance at the earth's surface. Almost all of it is in the oceans, which cover 70% of the surface area of the earth. However, the amounts of water present in the atmosphere and on land (as surface runoff, lakes and streams) is great enough to make it a significant agent in transporting substances between the lithosphere and the oceans.

Water interacts with both the atmosphere and the lithosphere, acquiring solutes from each, and thus provides the major chemical link between these two realms. The various transformations undergone by water through the different stages of the hydrologic cycle act to transport both dissolved and particulate substances between different geographic locations.

Table2.1.2.2.1**: Inventory of water on Earth**

Reservoir	Volume / 10^6, km^3	Percent of total
Oceans	1370	97.25
Ice caps and glaciers	39	2.05
Deep groundwater	5.3	0.38
Shallow groundwater	4.3	0.30
Lakes	0.125	0.01
Soil moisture	0.065	0.005
Atmosphere	0.013	0.001
Rivers	0.0017	0.0001
Biosphere	0.0006	0.00004
Total	**1408.7**	**100**

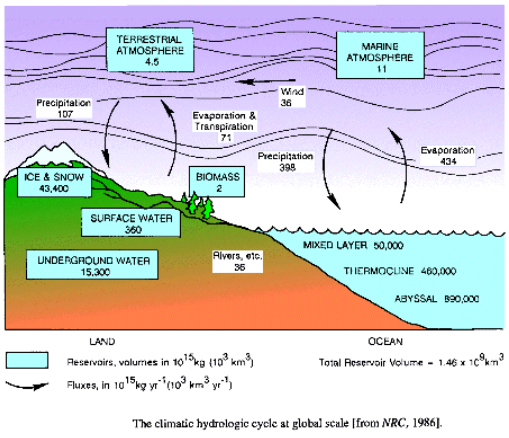
Where did the water come from?

It appears to have been bound up in the silica-based materials such as micas and amphiboles which accreted to form the Earth. The heat released during this process would have been sufficient to drive off this water, which amounted to about 0.01% by mass of the primordial material.

The hydrologic cycle

The *hydrologic cycle* refers to the steady state that exists between evaporation, condensation, percolation, runoff, and circulation of water. The cycle is driven by solar energy, mainly through direct vaporization, but also by convective motion induced by uneven heating.

The major interphase transport process of the hydrologic cycle is evaporation of water from the ocean. However, 90% of this vapor falls directly back into the ocean as rain, while 10% is transported over the land. Of the latter, about two-thirds evaporates again and one third runs off to the ocean.



The climatic hydrologic cycle at global scale [from *NRC*, 1986].

Figure 2.1.2.2.1: The climatic hydrological cycle at global scale (Diagram and text from a page at the Web site of the Institute for Global Environment and Society)

The movement of water on the earth's surface and through the atmosphere is known as the hydrologic cycle. Water is taken up by the atmosphere from the earth's surface in vapor form through evaporation. It may then be moved from place to place by the wind until it is condensed back to its liquid phase to form clouds. Water then returns to the surface of the earth in the form of either liquid (rain) or solid (snow, sleet, etc.) precipitation. Water transport can also take place on or below the earth's surface by flowing glaciers, rivers, and ground water flow.

The amounts of water precipitated onto the land and oceans are in approximate proportion to the relative surface areas, but evaporation from the ocean exceeds that from the land by about 37,400 km^{-3} yr^{-1}. This difference is the amount of water transported to the oceans by river runoff. When water condenses from the atmosphere in the form of rain, it is slightly enriched in H_2O^{18}. During epochs of glacial buildup the fraction of H_2O^{18} in the oceans consequently decreases. Observation of H_2O^{18}/H_2O^{16} ratios in marine sediments is one way of studying the timing and extent of past glaciations. Since the degree of heavy isotope enrichment of condensed water is temperature dependent, this same method can be used to estimate mean world temperatures in the distant past.

The hydrologic cycle also has important effects on the energy budget of the earth. Atmospheric water vapor (along with carbon dioxide and methane) tends to absorb the long-wavelength infrared radiation emitted by the earth's surface, partially trapping the incoming shorter-wavelength energy and thus maintaining the mean surface temperature about 30° higher than would be the case in the absence of water vapor. Of the 51% of the solar radiation incident on the atmosphere that reaches the earth's surface, about half of this (23%) is used to evaporate water. During the ten days that an average molecule resides in the atmosphere, it will travel about 1000 km. The atmospheric transport of water from equatorial to subtropical regions serves as an important mechanism for the transport of thermal energy; at latitudes of about 40°, as much as one-third of the energy input comes from release of latent heat from water vapor formed in equatorial regions.

Oceanic circulation

About 97% of the earth's water is contained in the two reservoirs which comprise the oceans. The upper mixed layer contains about 5% of the total; it is separated from the deeper and colder layer by the thermocline. Mixing between these two stratified layers is very slow; of the total ocean volume of 6.8 $\times 10^{18}$ m^3, only about 0.71 $\times 10^{15}$ m^3, or about 0.01%, moves between the two layers per year. The mean residence time of a water molecule in the deep layer is about 1600 years.

The large-scale motions of ocean water are the primary means by which chemical substances, especially those taken up and excreted by organisms, are transported within the ocean. An understanding of the general patterns of this circulation is essential in order to analyze the observed distribution of many of the chemical elements in different parts of the ocean and in the oceanic sediments.

Atmospheric circulation

The circulation of the surface waters of the ocean is driven by the prevailing winds. The latter arise from uneven heating of the earth's surface, and are arranged in bands that parallel the equator.

Although the motion of the waters at the surface of the ocean are driven by the winds, they do not follow them in a simple manner. The reasons are threefold: the Coriolis effect, the presence of land masses, and uneveness in the sea level due to regional differences in temperature and atmospheric pressure.

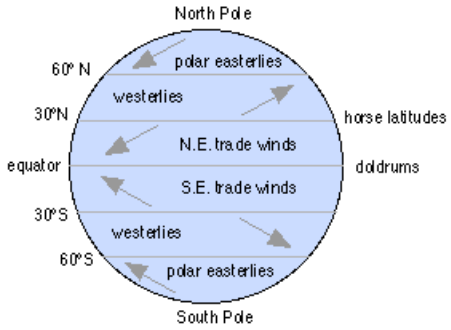


Figure 2.1.2.2.2: Atmospheric winds

The most intense heat input into the atmosphere occurs near the equator, where the heated air rises and cools, producing intense local precipitation but little surface wind. After cooling and losing moisture, this air moves north and south and descends at a latitude of about 30°. As it descends, it warms (largely by adiabatic compression) and its relative humidity decreases. The extreme dryness of this air gives rise to the subtropical desert regions between about 15° and 30°. Part of this air flows back toward the equator, giving rise to the northeast and southwest trade winds; the deflection to the east or west is caused by the Coriolis effect. Another part of the descending air travels poleward, producing the prevailing westerlies. Eventually these collide with cold air masses moving away from the polar regions, producing a region of unstable air and storm activity known as a polar front. Some of this polar air picks up enough heat to rise and enter into polar cell circulation patterns.

The flow of air in the prevailing westerlies is subject to considerable turbulence which gives rise to planetary waves. These are moving regions in which warm surface air is lifted to higher levels, producing lines of storms that travel from west to east, and exchanging more air between the polar and temperate regions.

Surface currents of the oceans

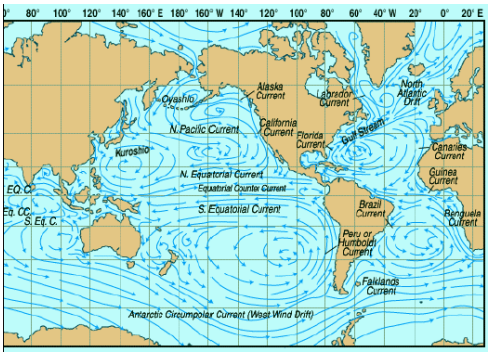


Figure 2.1.2.2.3:

In the Northern hemisphere, the Coriolis effect not only deflects south-moving objects to the east but it also causes currents flowing parallel to the equator to veer to the right of the direction of flow, i.e. to the north or south.

In addition, prevailing westerly winds and the eastward rotation of the earth cause water to pile up by a few centimeters at the western edges of the oceans. The resultant downhill flow, interacting with Coriolis forces, produces a *western boundary current* that runs south-to-north in the northern hemisphere. A similar but opposite effect gives rise to a south-flowing *eastern boundary current* on continental east coasts.

Thermohaline circulation

In contrast to the upper levels of the ocean, the deep ocean is *stratified*; the density increases with depth so as to inhibit the vertical transport of water. This stratification divides the deep oceans into several distinct water masses which undergo movement in a more or less horizonal plane, with adjacent masses sometimes moving in opposite directions.

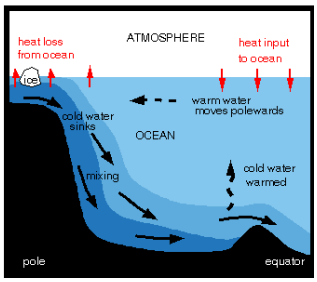


Figure 2.1.2.2.4: As the cold water fills up the deepest regions and spills over ridges into other deep basins, it creates huge undersea cascades which rival the greatest terrestrial waterfalls in height and the largest rivers in volume. For more on this, see the Feb. 1989 *Scientific American*

The winds and atmospheric effects outlined above affect only the upper part of the ocean. Below 100 meters or so, oceanic circulation is driven by the *density* of the seawater, which is determined by its temperature and its salinity. Variations in these two quantities give rise to the *thermohaline circulation* of the deep currents of the ocean.

It all starts when seasonal ice begins to form in the polar regions. Because the salts dissolved in seawater cannot be accommodated within the ice structure, they are largely excluded from the new ice and remain in solution. This increases the density of the surrounding unfrozen water, causing it to sink into the bottom ocean.

There are two major locations at which surface waters enter the deep ocean. The northern entry point is in the Norwegian Sea off Greenland; this water forms a mass known as the North Atlantic Deep Water (NADW) which flows southward across the equator.

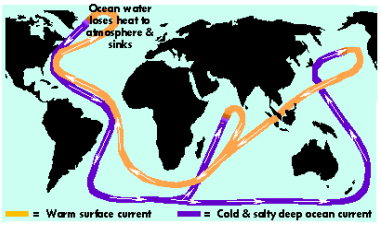
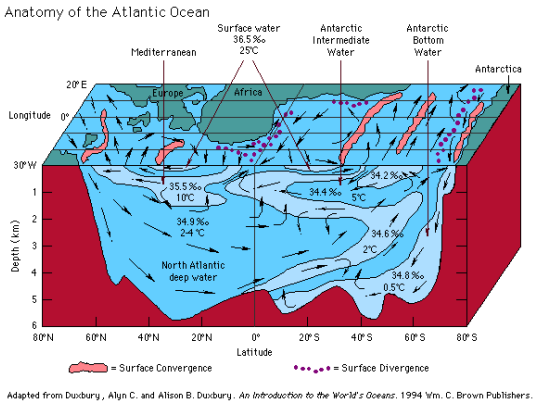


Figure 2.1.2.2.5: From Don Reed, Marine Geology 130 at San Jose State U.

Most of the transport into the deep ocean takes place in the Weddell Sea off the coast of Antarctica. The highly saline water flows down the submerged Antarctic Slope to begin a 5000-year trip to the north across the bottom of the ocean. This is the major route by which dissolved CO_2 and O_2 (which are more soluble in this cold water) are transported into the deep ocean where it forms a water mass known as the Antarctic Bottom Water (AABW) which can be traced into all three oceans.

In the south, a flow from the antarctic region forms a water mass known as the Antarctic Bottom Water (AABW) which can be traced into all three oceans.

The Pacific Ocean lacks any major identifiable direct source of cold water, so it is less differentiated and its deep circulation is sluggish and poorly defined.



Adapted from Duxbury, Alyn C. and Alison B. Duxbury. *An Introduction to the World's Oceans*. 1994 Wm. C. Brown Publishers.

Figure 2.1.2.2.6: Temperature and salinity profiles in a north-south subsection of the Atlantic ocean.

As is apparent from the figure above, the vertical profiles of temperature and especially of the salinity are not uniform. To some extent, these two parameters have opposite effects: in equatorial regions, temperatures are higher (leading to lower density) but evaporation rates are also higher (leading to higher density). In polar regions, the formation of sea ice raises the density of the seawater (because only a small proportion of salt is incorporated into the ice).

The nature and extent of the deep ocean currents differ in the Atlantic, Pacific, and Indian oceans. These currents are much slower than the surface currents, and in fact have not been measured directly; their existence is however clearly implied by the chemical composition and temperature of water samples taken from various parts of the ocean. Estimated rates are of the order of kilometers per month, in contrast to the few kilometers per hour of surface waters. The deep currents are the indirect results of processes occuring at the surface in which cold water of high salinity is produced as sea ice forms in the arctic and antarctic regions. This water is so dense that it sinks to the bottom, displacing warmer or less saline water as it moves.

Coastal Upwelling

Recirculation of deep water to the surface occurs to a very small extent in many regions, but it is especially pronounced where water entering the Antarctic Bottom mass displaces other bottom water, and where water piles up at the western edges of continents. This latter water flows downhill (forming the western boundary currents mentioned above) and is replaced by colder water from the deep ocean. The deep ocean contains few organisms to deplete the water of the nutrients it receives from the remains of the dead organisms floating down from above; this upwelled water is therefor exceptionally rich in nutrients, and strongly encourages the growth of new organisms that extend up the food chain to fish. Thus the wind-driven upwelling that occurs off the west coast of South America is responsible for the Peruvian fishing and guano fertilizer industry.

> *About every seven years these prevailing winds disappear for a while, allowing warm equatorial waters to move in. This phenomenon is known as El Niño and it results in massive kills of plankton and fish. Decomposition of the dead organisms reduces the oxygen content of the water, causing the death of still more fish, and allowing reduced compounds such as hydrogen sulfide to accumulate.*

Matthias Tomczak's Introduction to physical oceanography is an excellent source of more information on the topics covered above.

Page last modified: 21.01.2008

© 1998 by Stephen Lower

For information about this Web site or to contact the author,
please see the*Chem1 Virtual Textbook home page*.

SOME RIGHTS RESERVED

This work is licensed under a Creative Commons Attribution-NonCommercial 2.5 License.

This page titled 2.1.2.2: The hydrosphere and the oceans is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.2.3: Chemistry and geochemistry of the oceans

Composition of seawater

The composition of the ocean has attracted the attention of some of the more famous names in science, including Robert Boyle, Antoine Lavoisier and Edmund Halley. Their early investigations tended to be difficult to reproduce, owing to the different conditions under which they crystallized the various salts.

As many as 54 salts, double salts and hydrated salts can be obtained by evaporating seawater to dryness. At least 73 elements are now known to be present in seawater.

Major ions of seawater These values, expressed in parts per thousand, are for seawater of 35% salinity.

cations	g/kg	anions	g/kg
Na^+	10.77	Cl^-	19.354
Mg^{2+}	1.29	SO_4^{2-}	2.712
Ca^{2+}	0.412	Br^-	0.087
K^+	0.399		
Sr^{2+}	0.0079		
Al^{3+}	0.005		

The best way of characterizing seawater is in terms of its ionic content, shown above. The remarkable thing about seawater is the constancy of its relative ionic composition. The overall salt content, known as the salinity (grams of salts contained in 1 kg of seawater), varies slightly within the range of 32-37.5%, corresponding to a solution of about 0.7% salt content. The ratios of the concentrations of the different ions, however, are quite constant, so that a measurement of Cl^- concentration is sufficient to determine the overall composition and total salinity. Although most elements are found in seawater only at trace levels, marine organisms may selectively absorb them and make them more detectable. Iodine, for example, was discovered in marine algae (seaweeds) 14 years before it was found in seawater. Other elements that were not detected in seawater until after they were found in marine organisms include barium, cobalt, copper, lead, nickel, silver and zinc. Si^{32}, presumably deriving from cosmic ray bombardment of Ar, has been discovered in marine sponges.

pH balance. Reflecting this constant ionic composition is the pH, which is usually maintained in the narrow range of 7.8-8.2, compared with 1.5 to 11 for fresh water. The major buffering action derives from the carbonate system, although ion exchange between Na+ in the water and H^+ in clay sediments has recently been recognized to be a significant factor.

Conservative and non-conservative substances

The major ionic constituents whose concentrations can be determined from the salinity are known as conservative substances. Their constant relative concentrations are due to the large amounts of these species in the oceans in comparison to their small inputs from river flow. This is another way of saying that their residence times are very large.

Ion	Concentration in river water micromols/L	Concentration in sea water micromols/L	Residence time/10^6 years
Cl^-	250	558,000	87
Na^+	315	479,000	55
Mg^{2+}	150	54,300	13
SO_4^{2-}	120	28,900	8.7
Ca^{2+}	367	10,500	1
K^+	36	10,400	10
HCO_3^-	870	2000	0.083
H_4SiO_4	170	100	0.021
NO_3^-	10	20	0.072
$H_2PO_4^-$	0.7	1	0.080

Replacement time with respect to river addition for some components of seawater

A number of other species, mostly connected with biological activity, are subject to wide variations in concentration. These include the nutrients NO_3^-, NO_2^-, NH_4^+, and HPO_4^{2-}, which may become depleted near the surface in regions of warmth and light. As was explained in the preceding subsection on coastal upwelling, offshore prevailing winds tend to drive Western coastal surface waters out to sea, causing deeper and more nutrient-rich water to be drawn to the surface. This upwelled water can support a large population of phytoplankton and thus of zooplankton and fish. The best-known example of this is the anchovy fishery off the coast of Peru, but the phenomenon occurs to some extent on the West coasts of most continents, including our own.

Other non-conservative components include Ca^{2+} and SiO_4^{2-}. These ions are incorporated into the solid parts of marine organisms, which sink to greater depths after the organisms die. The silica gradually dissolves, since the water is everywhere undersaturated in this substance. Calcium carbonate dissolves at intermediate depths, but may reprecipitate in deep waters owing to the higher pressure. Thus the concentrations of Ca and of SiO_4^{2-} tend to vary with depth. The gases O_2 and CO_2, being intimately involved with biological activity, are also non-conservative, as are N_2O and CO.

Organic matter

Most of the organic carbon in seawater is present as dissolved material, with only about 1-2% in particulates. The total organic carbon content ranges between 0.5 mg/L in deep water to 1.5 mg/L near the surface. There is still considerable disagreement about the composition of the dissolved organic matter; much of it appears to be of high molecular weight, and may be polymeric. Substances qualitatively similar to the humic acids found in soils can be isolated. The greenish color that is often associated with coastal waters is due to a mixture of fluorescent, high molecular weight substances of undetermined composition known as *"Gelbstoffe"*. It is likely that the significance of the organic fraction of seawater may be much greater than its low abundance would suggest. For one thing, many of these substances are lipid-like and tend to adsorb onto surfaces. It has been shown that any particle entering the ocean is quickly coated with an organic surface film that may influence the rate and extent of its dissolution or decomposition. Certain inorganic ions may be strongly complexed by humic-like substances. The surface of the ocean is mostly covered with an organic film, only a few molecular layers thick. This is believed to consist of hydrocarbons, lipids, and the like, but glycoproteins and proteoglycans have been reported. If this film is carefully removed from a container of seawater, it will quickly be reconstituted. How significant this film is in its effects on gas exchange with the atmosphere is not known.

Regulation of ocean composition

The salinity of the ocean appears to have been about the same for at least the last 200 million years. There have been changes in the relative amounts of some species, however; the ratio of Na/K has increased from about 1:1 in ancient ocean sediments to its present

value of 28:1. Incorporation of calcium into sediments by the action of marine organisms has depleted the Ca/Mg ratio from 1:1 to 1:3.

Mass balance of P, C, and Ca for the oceans

Element	input to ocean	dissolved in seawater	in dead organisms	loss to sediments	residence time, y
phosphorus	1 (erosion)	1	1	1 (organic)	10,000
carbon	100 as CO_2 500 as carbonate	1000	125	100 (organic) 600 (carbonate)	165,000
calcium	500 (erosion)	5000	25	500 (as $CaCO_3$)	10^6

If the composition of the ocean has remained relatively unchanged with time, the continual addition of new mineral substances by the rivers and other sources must be exactly balanced by their removal as sediment, possibly passing through one or more biological systems in the process.

Where does the salt come from?

In 1715 Edmund Halley suggested that the age of the ocean (and thus presumably of the world) might be estimated from the rate of salt transport by rivers. When this measurement was actually carried out in 1899, it gave an age of only 90 million years.

> *This is somewhat better than the calculation made in 1654 by James Ussher, the Anglican Archbishop of Armagh, Ireland, based on his interpretation of the Biblical book of Genesis, that the world was created at 9 A.M. on October 23, 4004 BC, but it is still far too recent, being about when the dinosaurs became extinct. What Halley actually described was the residence time, which is about right for Na but much to long for some of the minor elements of seawater.*

The commonly stated view that the salt content of the oceans derives from surface runoff that contains the products of weathering and soil leaching is not consistent with the known compositions of the major river waters (See Table). The halide ions are particularly over-represented in seawater, compared to fresh water. These were once referred to as “excess volatiles”, and were attributed to volcanic emissions. With the discovery of plate tectonics, it became apparent that the locations of seafloor spreading at which fresh basalt flows up into the ocean from the mantle are also sources of mineral-laden water. Some of this may be seawater that has cycled through a hot porous region and has been able to dissolve some of the mineral material owing to the high temperature. Much of the water, however, is “juvenile” water that was previously incorporated into the mantle material and has never before been in the liquid phase. The substances introduced by this means (and by volcanic activity) are just the elements that are “missing” from river waters. Estimates of what fraction of the total volume of the oceans is due to juvenile water (most of it added in the early stages of mantle differentiation that began a billion years ago) range from 30 to 90%.

Geochemical processes involving the oceans

The oceans can be regarded as a product of a giant acid-base titration in which the carbonic acid present in rain reacts with the basic materials of the lithosphere. The juvenile water introduced at locations of ocean-floor spreading is also acidic, and is partly neutralized by the basic components of the basalt with which it reacts. Surface rocks mostly contain aluminum, silicon and oxygen combined with alkali and alkaline-earth metals, mainly potassium, sodium and calcium. The CO_2 and volcanic gases in rainwater react with this material to form a solution of the metal ion and HCO_3^-, in which is suspended some hydrated SiO^2. The solid material left behind is a clay such as kaolinite, $Al_2Si_2O_5(OH)_4$. This first forms as a friable coating on the surface of the weathered rock; later it becomes a soil material, then an alluvial deposit, and finally it may reach the sea as a suspended sediment. Here it may

undergo a number of poorly-understood transformations to other clay sediments such as illites. Sea floor spreading eventually transports these sediments to a subduction region under a continental block, where the high temperatures and pressures permit reactions that transform it into hard rock such as granite, thus completing the geochemical cycle.

Deep-sea hydrothermal vents are now recognized to be another significant route for both the addition and removal of ionic substances from seawater.

Distribution and cycling of elements in the oceans

Although the relative concentrations of most of the elements in seawater are constant throughout the oceans, there are certain elements that tend to have highly uneven distributions vertically, and to a lesser extent horizontally. Neglecting the highly localized effects of undersea springs and volcanic vents, these variations are direct results of the removal of these elements from seawater by organisms; if the sea were sterile, its chemical composition would be almost uniform.

Plant life can exist only in the upper part of the ocean where there is sufficient light available to drive photosynthesis. These plants, together with the animals that consume them, extract nutrients from the water, reducing the concentrations of certain elements in the upper part of the sea. When these organisms die, they fall toward the lower depths of the ocean as particulate material. On the way down, some of the softer particles, deriving from tissue, may be consumed by other animals and recycled. Eventually, however, the nutrient elements that were incorporated into organisms in the upper part of the ocean will end up in the colder, dark, and essentially lifeless lower part.

Mixing between the upper and lower reservoirs of the ocean is quite slow, owing to the higher density of the colder water; the average residence time of a water molecule in the lower reservoir is about 1600 years. Since the volume of the upper reservoir is only about 1/20 of that of the lower, a water molecule stays in the upper reservoir for only about 80 years.

Except for dissolved oxygen, all elements required by living organisms are depleted in the upper part of the ocean with respect to the lower part. In the case of the major nutrients P, N and Si, the degree of depletion is sufficiently complete (around 95%) to limit the growth of organisms at the surface. These three elements are said to be biolimiting. A few other biointermediate elements show partial depletion in surface waters: Ca (1%), C (15%), Ba (75%).

The organic component of plants and animals has the average composition $C_{80}N_{15}P$. It is remarkable that the ratio of N:P in seawater (both surface and deep) is also 15:1; this raises the interesting question of to what extent the ocean and life have co-evolved.

In the deep part of the ocean the elemental ratio corresponds to $C_{800}N_{15}P$, but of course with much larger absolute amounts of these elements. Eventually some of this deeper water returns to the surface where the N and P are quickly taken up by plants. But since plants can only utilize 80 out of every 800 carbon atoms, 90 percent of the carbon will remain in dissolved form, mostly as HCO_3^-.

> *To work out the balance of Ca and Si used in the hard parts of organisms, we add these elements to the average composition of the lower reservior to get $Ca_{3200}Si_{50}C_{800}N_{15}P$. Particulate carbon falls into the deep ocean in the ratio of about two atoms in organic tissue to one atom in the form of calcite. This makes the overall composition of detrital material something like $C_{120}N_{15}P$; i.e., 80 organic C's and 40 in $CaCO_3$. Accompanying these 40 calcite units will be 40 Ca atoms, but this represents a minor depletion of the 3200 Ca atoms that eventually return to the surface, so this element is only slightly depleted in the upper waters. Silicon, being far less abundant, is depleted to a much greater extent.*

The oceanic sediments

The particulate shower.

A continual rain of particulate material from dead organisms falls through the ocean. This shower is comprised of three major kinds of material: calcite ($CaCO_3$), silica (SiO_2), and organic matter. The first two come from the hard parts of both plants and animals (mainly microscopic animals such as foraminifera and radiolarians). The organic matter is derived mainly from the soft tissues of

organisms, and from animal fecal material. Some of this solid material dissolves before it reaches the ocean floor, but not usually before it enters the deep ocean where it will remain for about 1600 years.

The remainder of this material settles onto the floor of the sea, where it forms one component of a layer of sediments that provide important information about the evolution of the sea and of the earth. Over a short time scale of months to years, these sediments are in quasi-equilibrium with the seawater. On a scale of millions of years, the sediments are merely way-stations in the geochemical cycling of material between the earth's surface and its interior.

The oceanic sediments have three main origins:

- **Detrital** material is derived largely from particles deposited in the ocean by rivers and also directly by the wind. These materials are mostly aluminosilicates (clays), along with some quartz. These substances accumulate on the deep ocean floor at a rate of about 0.3 g cm^{-2} per 1000 years.
- **Authigenic** materials are formed by precipitation within the ocean or by crystallization within the sediment itself. These constitute only a small fraction of the total sediment.
- **Biogenic** components consist mainly of the calcium carbonate and silica that make up the hard parts of organisms. Scavenging by other organisms is so efficient that only about 0.3 % of the total deep sea sediment consists of actual organic material.

Our main interest lies with the silica and calcium carbonate, since these substances form a crucial part of the biogeological cycle. Also, their distributions in the ocean are not uniform- a fact that must tell us something.

The skeletons of diatoms and radiolarians are the principal sources of silica sediments. Since the ocean is everywhere undersaturated with respect to silica, only the most resistant parts of these skeletons reach the bottom of the deep ocean and get incorporated into sediments. Silica sediments are less common in the Atlantic ocean, owing to the lower content of dissolved silica.

The parts of the ocean where these sediments are increasing most rapidly correspond to regions of upwelling, where deep water that is rich in dissolved silica rises to the surface where the silica is rapidly fixed by organisms. Where upwelling is absent, the growth of the organisms is limited, and little silica is precipitated. Since deep waters tend to flow from the Atlantic into the Pacific ocean where most of the upwelling occurs, Atlantic waters are depleted in silica, and silica sediments are not commonly found in this ocean.

For calcium carbonate, the situation is quite different. In the first place, surface waters are everywhere supersaturated with respect to both calcite and aragonite, the two common crystal forms of $CaCO_3$. Secondly, Ca^{2+} and HCO_3^- are never limiting factors in the growth of the coccoliths (plants) and forams (animals) that precipitate $CaCO_3$; their production depends on the availability of phosphate and nitrogen. Because these elements are efficiently recycled before they fall into the deep ocean, their supply does not depend on upwelling, and so the production of solid is more uniformly distributed over the world's oceans.

More importantly, however, the chances that a piece of carbonate skeleton will end up as sediment will be highly dependent on both the local CO_3^{2-} concentration and the depth of the ocean floor. These factors give rise to small-scale variations in the production of carbonate sediments that can be quite wide-ranging.

Ocean sediments and continental drift

New crust is being generated and moving away from the crests of the mid-ocean ridges at a rate of a few centimetres per year. Although the crests of these ridges are relatively high points, projecting to within about 3000 m of the surface, the continual injection of new material prevents sediments from accumulating in these areas. Farther from the crests, carbonate sediments do build up, eventually reaching a depth of about 500 m, but by this time the elevation has dropped off below the saturation horizon, so from this point on the carbonate sediments are overlaid by red clay.

If we drill a hole down through a part of the ocean floor that is presently below the saturation horizon, the top part of the drill core will consist of clay, followed by $CaCO_3$ at greater depths. The core may also contain regions in which silica predominates. Since silica production is very high in equatorial regions, the appearance of such a layer suggests that this particular region of the oceanic crust has moved across the equator.

Page last modified: 21.01.2008

© 1998 by Stephen Lower

For information about this Web site or to contact the author,
please see the *Chem1 Virtual Textbook home page*.



This work is licensed under a Creative Commons Attribution-NonCommercial 2.5 License.

This page titled 2.1.2.3: Chemistry and geochemistry of the oceans is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.2.4: Chemical budgets of oceanic elements

To the extent that the composition of the ocean remains constant, the rate at which any one element is introduced into seawater must equal the rate of its removal. A listing of the various routes of addition and removal, together with the estimated rate of each process, constitutes the budget for a given element. If that budget is greatly out of balance and no other transport routes are apparent, then it is likely that the ocean is not in a steady state with respect to that element, at least on a short time scale. It is important to understand, however, that short-term deviations from constant composition are not necessarily inconsistent with a long-term steady state. Deviations from the latter condition are most commonly inferred from geological evidence.

The major input of elements to the oceans is river water. Groundwater seepage constitutes a very small secondary source. These were considered the only sources until the 1970's, when the existence of hydrothermal springs at sites of seafloor spreading became known. There are presently no reliable estimates of the magnitude of this source. Pollution represents an additional input, mainly dissolved in river water, but also sometimes in rain and by dry deposition.

Routes of removal are formation and burial of sediments, formation of evaporite deposits, direct input to the atmosphere by sea-salt aerosol transfer associated with bubble-breaking, and burial with sediments, either in interstitial water or adorbed onto active surfaces. Reaction with newly-formed basalt associated with undersea volcanic activity appears to be an important removal mechanism for some elements.

Table 7: Input and output rates of major elements in seawater by two routes. The units are Tg y–1. The other major routes involve interaction with the seafloor through hyrothermal vents and incorporation into sediments, but only river input and loss to the atmosphere are quantifiable with any degree of accuracy.

Species	Addition from Rivers	Loss to Atmosphere
Cl^-	308	40
Na^+	269	21
S (mostly SO_4^{2-})	143	4
Mg^{2+}	137	3
K^+	52	1
Ca^{2+}	550	0.5
HCO^{3-}	1980	
H_4SiO_4	180	

Phosphorus

The major elements undergoing steady-state dynamic change in the ocean are connected with biological processes. The key limiting element in the development of oceanic biomass is phosphorus, in the form of the phosphate ion. For terrestrial plant life, nitrogen is more commonly the limiting element, where it is taken up in the form of the nitrate ion. In the ocean, however, the ratio of the nitrate ion concentration to that of phosphorus has been found to be everywhere the same; this implies that the concentration of one controls that of the other. The source of nitrate ion is atmospheric N_2, which is freely soluble in water and is thus always present in abundance. The conversion of N_2 to NO_3^- is presumed to be biologically mediated, probably by bacteria. The constancy of the NO_3^-/P concentration ratio implies that the phosphorus concentration controls the activity of the nitrogen-fixing organisms, and thus the availability of nitrogen to oceanic life.

Photosynthetic activity in the upper part of the ocean causes inorganic phosphate to be incorporated into biomass, reducing the concentration of phosphorus; in warm surface waters, phosphate may become totally depleted. A given phosphorus atom may be traded several times among the plant, animal and bacterial populations before it eventually finds itself in biodebris (a dead organism or a fecal pellet) that falls into the deep part of the sea. Only about 1% of the phosphorus atoms that descend into deeper waters actually reach the bottom, where they are incorporated into sediments and permanently removed from circulation. The other 99% are released in the form of soluble phosphate, which is eventually brought to the surface in regions of upwelling. An average

phosphorus atom will undergo one cycle of this circulation in about 1000 years; only a few months of this cycle will be spent in biomass. After an average of 100 such cycles, the atom will be removed from circulation and locked into the bottom sediment, and a new one will have entered the sea with river or juvenile water.

Phosphorus is unique in that its major source of input to the oceans derives ultimately from pollution; in the long term, this represents a transfer of land-based phosphate deposits to the oceans. About half of the phosphorus input is in the form of suspended material, both organic and inorganic, the latter being in a variety of forms including phosphates adsorbed onto clays and iron oxide paticles, and calcium phosphate (apatatite) eroded from rocks. These various particulates are known to dissolve to some extent once they reach the ocean, but there is considerable uncertainty about the rates of these processes under various conditions.

The major sink for oceanic phosphorus is burial with organic matter; this accounts for about two-thirds of the phosphorus removed. Most of the remainder is due to deposition with $CaCO_3$. A minor removal route is through reaction with Fe(II) formed when seawater attacks hot basalt, and in the formation of evaporite deposits. However, there is more phosphorus in evaporite deposits in the Western U.S. than in all of the ocean, so it is apparent that the long-term phosphorus budget is still not clearly understood.

Carbon

Carbon enters the ocean from both the atmosphere (as CO_2) and river water, in which the principal species is HCO_3^-. Once in solution, the carbonate species are in equilibrium with each other and with H_3O^+, and the concentrations of all of these are influenced by the partial pressure of atmospheric CO_2. The mass budget for calcium is linked to that of carbon through solubility equilibria with the various solid forms of $CaCO_3$ (mainly calcite). During photosynthesis, C^{12} is taken up slightly more readily than the rare isotope C^{13}. Since the rate of photosynthesis is controlled by the phosphate concentration, the C^{13}/C^{12} ratio in the dissolved carbon dioxide of surface waters is slightly higher than in the ocean as a whole. Observations of carbon isotope ratios in buried sediments have been useful in tracking historical changes in phosphate concentrations.

The C:P ratio and the regulation of atmospheric CO_2

The ratio of carbon to phosphorus in sea salt is about eight times greater than the same ratio measured in organic debris. This implies that in exhausting the available phosphate, the living organisms in the upper ocean consume only 12.5% of the dissolved carbon. Even this relatively small withdrawal of carbon from the carbonate system is sufficient to noticeably reduce the partial pressure of gaseous CO_2 in equilibrium with the ocean; it has been estimated that if all life in the ocean should suddenly cease, the atmospheric CO_2 content would rise to about three times its present level. The regulation of atmospheric CO_2 pressure by the oceans also works the other way: since the amount of dissolved carbonate in the oceans is so much greater than the amount of CO_2 in the atmosphere, the oceans act to buffer the effects of additions of CO_2 to the atmosphere. Calculations indicate that about half of the CO_2 that has been produced by burning fossil fuel since the Industrial Revolution has ended up in the oceans.

Bicarbonate

This is of course the major carbonate species in the ocean. Although it is interconvertible with CO_2 and is thus coupled to the carbon and photosynthetic cycles, itself can neither be taken up nor produced by organisms, and so it can be treated somewhat independently of biological activity. In this sense the only major input of into the oceans is river water. The two removal mechanisms are formation of CO_2

$$H^+ + HCO_3^- \rightarrow H_2O + CO^2 \qquad (2.1.2.4.1)$$

and the (biologically mediated) formation of $CaCO_3^.$

$$Ca^{2+} + HCO_3^- \rightarrow CaCO_3 + H^+ \qquad (2.1.2.4.2)$$

Since the pH of the oceans does not change, H^+ is conserved and the removal of HCO_3^- by biogenic secretion of $CaCO_3$ can be expressed by the sum of these reactions:

$$Ca^{2+} + 2HCO_3^- \rightarrow CaCO_3 + CO_2 + H_2O \qquad (2.1.2.4.3)$$

whose reverse direction represents the dissolution of the skeletal remains of dead organisms as they fall to lower depths.

Calcite

The upper parts of the ocean tend to be supersaturated in $CaCO_3$. Solid $CaCO_3$, in the form of calcite, is manufactured by a large variety of organisms such as foraminifera. A constant rain of calcite falls through the ocean as these organisms die. The solubility of increases with pressure, so only that portion of the calcite that falls to shallow regions of the ocean floor is incorporated into sediments and removed from circulation; the remainder dissolves after reaching a depth known as the lysolcline. At the present time, the amounts of carbonate and Ca^{2+} supplied by erosion and volcanism appear to be only about one-third as great as the amount of calcite produced by organisms. As the carbonate concentration in a given region of the ocean becomes depleted due to higher calcite production, the lysocline moves up, tending to replenish the carbonate, and reducing the amount that is withdrawn by burial in sediments. Organic residues that fall into the deep sea are mostly oxidized to CO_2, presumably by bacterial activity.

Other Elements

Calcium

Calcium is removed from seawater solely by biodeposition as $CaCO_3$, a process whose rate can be determined quite accurately both at the present and in the past.

Table 2.1.2.4.1**: Present-day budget for oceanic calcium (Tg Ca/yr).**

inputs		**outputs**	
rivers	550	$CaCO_3$ into shallow water	520
volcanic basalt	191	$CaCO_3$ into deep water	440
cation exchange	37		
total	**778**		**960**

As is explained above, the upper part of the ocean is supersaturated in calcite but the lower ocean is not. For this reason, less than 20% of the $CaCO_3$ that is formed ends up as sediment and is eventually buried.

The main questions about the calcium budget tend to focus on the rates and locales at which dissolution of skeletal carbonates occurs, and on how to interpret the various kinds of existing carbonate sediments. For example, the crystalline form aragonite is less stable than calcite, and will presumably dissolve at a higher elevation. The absence of aragonite-containing pteropod shells in deeper deposits seems to confirm this, but in the absence of rate data is it difficult to know at what elevations these particular organisms originated.

The data in the above Table indicate that at the present time there is a net removal of calcium from the oceans. This is due to the rise in sea level since the decline of the more recent glacial epoch during the past 11,000 years. The additional water has covered the continental shelves, increasing the amount of shallow ocean where the growth of organisms is most intense. Over the more distant past (25 million years) the calcium budget appears to be well balanced.

Chloride

Evidence from geology and peleontology indicates that the salinity, and hence the chloride concentration of sewater has been quite constant for about 600 million years. There have been periods when climatic conditions and coastal topography have led to episodes of evaporite formation, but these have evidently been largely compensated by the eventual return of the evaporite deposits to the sea. The natural input of chloride from rivers is about 215 Tg/yr, but the present input is about half again as great (Table 7 on page 33), due to pollution. Also, there are presently no significant areas where seawater is evaporating to dryness. Thus the oceanic chloride budget is considerably out of balance. However, the replacement time of Cl^- in the oceans is so long (87 million years) that this will probably have no long-term effect.

Sodium

Although sodium is tied to chloride, it is also involved in the formation of silicate minerals, the weathering of rocks, and in cation exchange with clay sediments. Its short-term budget is quite out of balance for the same reasons as is that of chloride. On a longer

time scale, removal of sodium by reaction with hot basalt associated with undersea volcanic activity may be of importance.

Sulfate

Considerably more sulfate is being added to seawater than is being removed by the major mechanisms of sediment formation (mainly $CaSO_4$ and pyrites). The natural river input is 82 Tg of S per year, while that due to pollution is 61 Tg/yr from rivers and 17 Tg/yr from rain and dry deposition.

Magnesium

This element is unusual in that its river-water input is balanced mostly by reaction with volcanic basalt; removal through biogenic formation of magnesian calcite (dolomite) accounts for only 11% of its total removal from the ocean. The present-day magnesium budget seems to have been balanced for the past 100 million years. However, most of the extensive dolomite deposits were formed prior to this time, so the longer-term magnesium budget is poorly understood.

Potassium

The potassium budget of the ocean is not well understood. The element is unusual in that only about 60% of its input is by rivers; the remainder is believed to come from newly formed undersea basalt. The big question about potassium is how it is removed; fixation by ion-exchange with illite clays seems to be a major mechanism, and its uptake by basalt (at lower temperatures than are required for its release) is also believed to occur.

Silica

About 85% of the silicon input to the oceans comes from river water in the form of silicic acid, H_4SiO_4. The remainder probably comes from basalt. It is removed by biogenic deposition as opaline silica, $SiO_2 \cdot n\ H_2O$ produced mainly by planktonic organisms (radiolaria and diatoms). Unlike the case for $CaCO_3$, the ocean is everywhere undersaturated in silica, especially near the surface where these organisms deplete it with greater efficiency than any other element.

Because opaline silica dissolves so rapidly, only a small fraction makes it to the bottom. The major deposits occur in shallower waters where coastal upwelling provides a good supply of N and P nutrients for siliceous organisms. Thus over half of the biogenic silica deposits are found in the Antarctic ocean.

In spite of the fact that dissolved silica has the shortest replacement time (21,000 years) of any major element in the ocean, its concentration appears to have been remarkably constant during geological time. This is taken as an indication of the ability of siliceous organisms to respond quickly to changes in local concentrations of dissolved silica.

Nitrogen

This element is complicated by its biologically-mediated exchange with atmospheric nitrogen, and by its existence in several oxidation states, all of which are interconvertible. Unlike the other major elements, nitrogen does not form extensive sedimentary deposits; most of the nitrogen present in dead organic material seems to be removed before it can be buried. Through this mechanism there is extensive cycling of nitrogen between the shallow and deep parts of the oceans.

The real difficulty in constructing a budget for oceanic nitrogen is the very large uncertainty in the rates of the major input (fixation) and output. Both of these processes are biologically mediated, but little is known about what organisms are responsible, where they thrive and how they are affected by local nutrient supply and other conditions.

Contributors and Attributions

- Stephen Lower, Professor Emeritus (Simon Fraser U.) Chem1 Virtual Textbook

This page titled 2.1.2.4: Chemical budgets of oceanic elements is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

SECTION OVERVIEW

2.1.3: The Atmosphere

Life as we know it on the Earth is entirely dependent on the tenuous layer of gas that clings to the surface of the globe, adding about 1% to its diameter and an insignificant amount to its total mass. And yet the atmosphere serves as the earth's window and protective shield, as a medium for the transport of heat and water, and as source and sink for exchange of carbon, oxygen, and nitrogen with the biosphere. The atmosphere acts as a compressible fluid tied to the earth by gravitation; as a receptor of solar energy and a thermal reservoir, it constitutes the working fluid of a heat engine that transports and redistributes matter and energy over the entire globe. The atmosphere is also a major temporary repository of a number of chemical elements that move in a cyclic manner between the hydrosphere, atmosphere, and the upper lithosphere. Finally, the atmosphere is a site for a large variety of complex photochemically initiated reactions involving both natural and anthropogenic substances.

2.1.3.1: Structure and Composition of the Atmosphere

2.1.3.2: Origin and Evolution of the Atmosphere

This page titled 2.1.3: The Atmosphere is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.3.1: Structure and Composition of the Atmosphere

Life as we know it on the Earth is entirely dependent on the tenuous layer of gas that clings to the surface of the globe, adding about 1% to its diameter and an insignificant amount to its total mass. And yet the atmosphere serves as the earth's window and protective shield, as a medium for the transport of heat and water, and as source and sink for exchange of carbon, oxygen, and nitrogen with the biosphere. The atmosphere acts as a compressible fluid tied to the earth by gravitation; as a receptor of solar energy and a thermal reservoir, it constitutes the working fluid of a heat engine that transports and redistributes matter and energy over the entire globe. The atmosphere is also a major temporary repository of a number of chemical elements that move in a cyclic manner between the hydrosphere, atmosphere, and the upper lithosphere. Finally, the atmosphere is a site for a large variety of complex photochemically initiated reactions involving both natural and anthropogenic substances.

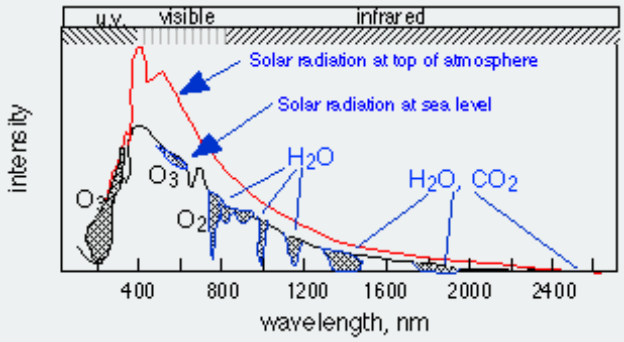
> *On the scale of cubic meters the air is a homogeneous mixture of its constituent gases, but on a larger scale the atmosphere is anything but uniform. Variations of temperature, pressure, and moisture content in the layers of air near the earth's surface give rise to the dynamic effects we know as the weather.*

Although the density of the atmosphere decreases without limit with increasing height, for most practical purposes one can roughly place its upper boundary at about 500 km. However, half the mass of the atmosphere lies within 5 km, and 99.99% within 80 km of the surface. The average atmospheric pressure at sea level is 1.01×10^5 pascals, or 1010 millibars. A 1-cm^2 cross section of the earth's surface supports a column weighing 1030 g; the total mass of the atmosphere is about 5.27×10^{21} g.

About 80% of the mass of the atmosphere resides in the first 10 km; this well-mixed region of fairly uniform composition is known as the *troposphere*.

Solar irradiation of the Earth

The gases ozone, water vapor, and carbon dioxide are only minor components of the atmosphere, but they exert a huge effect on the Earth by absorbing radiation in the ranges indicated by the shading. Ozone in the upper atmosphere filters out the ultraviolet light below about 360 nm that is destructive of life. O_2, H_2O, CO_2 and CH_4 are "greenhouse" gases that trap some of the heat absorbed from the Sun and prevent it from re-radiating back into space.

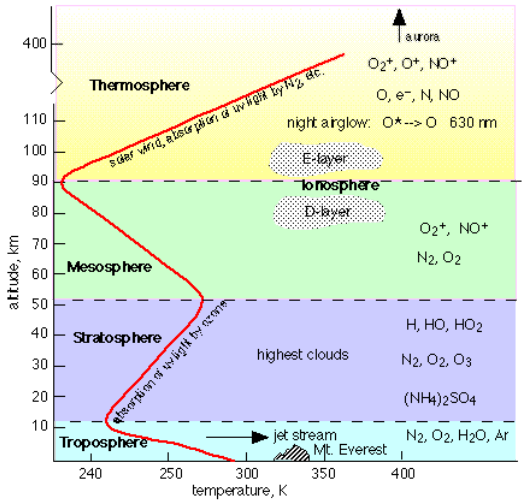


Structure of the atmosphere

We commonly think of gas molecules as moving about in a completely random manner, but the Earth's gravitational field causes downward motions to be very slightly favored so that the molecules in any thin layer of the air collide more frequently with those in the layer below. This gives rise to a pressure gradient that is the most predictable and well known structural characteristic of the atmosphere. This gradient is described by an exponential law which predicts that the atmospheric pressure should decrease by 50% for every 6 km increase in altitude. This law also predicts that the composition of a gas mixture will change with altitude, the lower-molecular weight components being increasingly favored at higher altitudes. However, this gravitational fractionation effect is completely obliterated below about 160 km owing to turbulence and convective flows (winds).

The atmosphere is divided vertically into several major regions which are distinguished by the sign of the temperature gradient. In the lowermost region, the troposphere, the temperature falls with increasing altitude. The major source of heat input into this part of the atmosphere is long-wave radiation from the earth's surface, while the major loss is radiation into space.

At higher elevations the temperature begins to rise with altitude as we move into a region in which heat is produced by exothermic chemical reactions, mainly the decomposition of ozone that is formed photochemically from dioxygen in the stratosphere. At still higher elevations the ozone gives out and the temperature begins to drop; this is the mesosphere, which is finally replaced by the thermosphere which consists largely of a plasma (gaseous ions). This outer section of the atmosphere which extends indefinitely to perhaps 2000 km is heated by absorption of intense u.v. radiation from the Sun and also from the solar wind, a continual rain of electrons, protons, and other particles emitted from the Sun's surface.



Structure of the atmosphere. *The main divisions of the atmosphere are defined by the elevations at which the sign of the temperature gradient changes. The chemical formulas at the right show the major species of interest in the various regions. The shaded D- and E-layers are regions of high ion concentrations that reflect radio waves and are important in long-distance communication.*

Composition of the atmosphere

Except for water vapor, whose atmospheric abundance varies from practically zero up to 4%, the fractions of the major atmospheric components N_2, O_2, and Ar are remarkably uniform below about 100 km. At greater heights, diffusion becomes the principal transport process, and the lighter gases become relatively more abundant. In addition, photochemical processes result in the formation of new species whose high reactivities would preclude their existence in significant concentrations at the higher pressures found at lower elevations.

The atmospheric gases fall into three abundance categories: major, minor, and trace. Nitrogen, the most abundant component, has accumulated over time as a result of its geochemical inertness; a very small fraction of it passes into the other phases as a result of biological activity and natural fixation by lightning. It is believed that denitrifying bacteria in marine sediments may provide the major route for the return of N_2 to the atmosphere. Oxygen is almost entirely of biological origin, and cycles through the hydrosphere, the biosphere, and sedimentary rocks. Argon consists mainly of Ar^{40} which is a decay product of K^{40} in the mantle and crust.

Table 1: Major components

nitrogen	N_2	78.08 %
oxygen	O_2	20.95 %
argon	Ar	0.93 %

The most abundant of the minor gases aside from water vapor is carbon dioxide, about which more will be said below. Next in abundance are neon and helium. Helium is a decay product of radioactive elements in the earth, but neon and the other inert gases are primordial, and have probably been present in their present relative abundances since the earth's formation. Two of the minor gases, ozone and carbon monoxide, have abundances that vary with time and location. A variable abundance implies an imbalance between the rates of formation and removal. In the case of carbon monoxide, whose major source is anthropogenic (a small amount is produced by biological action), the variance is probably due largely to localized differences in fuel consumption, particularly in internal combustion engines. The nature of the carbon monoxide sink (removal mechanism) is not entirely clear; it may be partly microbial.

Table 2: Minor components

water	H_2O	0-4 %
carbon dioxide	CO_2	325 ppm
neon	Ne	18 ppm
helium	He	5 ppm
methane	CH_4	2 ppm
krypton	Kr	1 ppm
hydrogen	H_2	0.5 ppm
nitrous oxide	N_2O	0.3 ppm
carbon monoxide	CO	0.05-0.2 ppm
ozone	O_3	0.02 - 10 ppm
xenon	Xe	0.08 ppm

Table 3:Trace components

ammonia	NH_3	4 ppb
nitrogen oxide	NO	1 ppb
sulfur dioxide	SO_2	1 ppb
hydrogen sulfide	H_2S	0.05 ppb

Ozone

Ozone is formed by the reaction of O_2 with oxygen atoms produced photochemically. As a consequence the abundance of ozone varies with the time of day, the concentration of O atoms from other sources (photochemical smog, for example), and particularly with altitude; at 30 km, the ozone concentration reaches a maximum of 12 ppm.

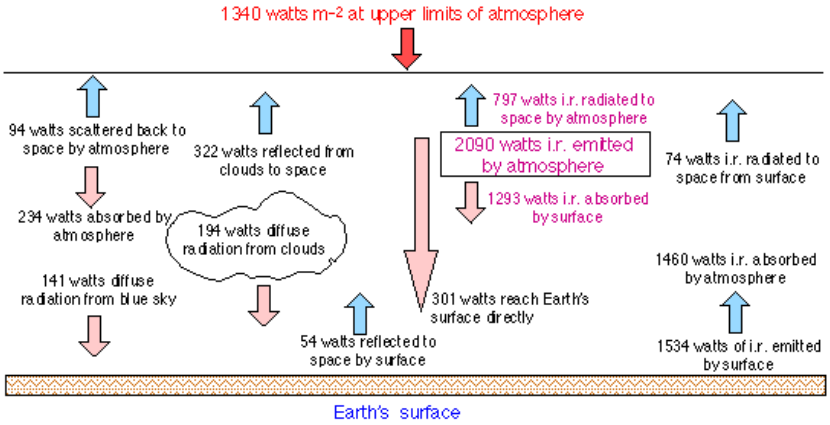
Carbon dioxide

The concentration of atmospheric carbon dioxide, while fairly uniform globally, is increasing at a rate of 0.2-0.7% per year as a result of fossil fuel burning. The present CO_2 content of the atmosphere is about 129 10^{18} g. Most of the CO^2, however, is of natural origin, and represents the smallest part of the total carbonate reservoir that includes oceanic CO_2, HCO_3^-, and carbonate sediments. The latter contain about 600 times as much CO_2 as the atmosphere, and the oceans contain about 50 times as much. These relative amounts are controlled by the rates of the reactions that interconvert the various forms of carbonate.

The surface conditions on the earth are sensitively dependent on the atmospheric CO_2 concentration. This is due mainly to the strong infrared absorption of CO_2, which promotes the absorption and trapping of solar heat (see below). Since CO_2 acts as an acid in aqueous solution, the pH of the oceans is also dependent on the concentration of CO_2 in the atmosphere; it has been estimated that if only 1% of the carbonate presently in sediments were still in the atmosphere, the pH of the oceans would be 5.9, instead of the present 8.2.

Energy balance of the atmosphere and earth

The amount of energy (the solar flux) impinging on the outer part of the atmosphere is 1367 watts m^{-2} . About 30% of this is reflected or scattered back into space by clouds, dust, and the atmospheric gas molecules themselves, and by the earth's surface. About 19% of the radiation is absorbed by clouds or the atmosphere (mainly by and O_3 , but not CO_2), leaving 51% of the incident energy available for absorption by the earth's surface. If one takes into account the uneven illumination of the earth's surface and the small flux of internal heat to the surface, the assumption of thermal equilibrium requires that the earth emit about 240 watts m^{-2}. This corresponds to the power that would be emitted by a black body at 255 K, or –18°C, which is the average temperature of the atmosphere at an altitude of 5 km. The observed mean global surface temperature of the earth is 13°C, and is presumably the temperature required to maintain thermal equilibrium between the earth and the atmosphere.



The greenhouse effect

The energy radiated by the earth has a longer wavelength (maximum 12m) than the incident radiation. Most gases absorb radiation in this range quite efficiently, including those gases such as CO_2 and N_2O that do not absorb the incident radiation. The energy absorbed by atmospheric gases is re-radiated in all directions; some of it therefore escapes into space, but a portion returns to the earth and is reabsorbed, thus raising its temperature.This is commonly called the greenhouse effect. If the amount of an infrared-absorbing gas such as carbon dioxide increases, a larger fraction of the incident solar radiation is trapped, and the mean temperature of the earth will increase.

Any significant increase in the temperature of the oceans would increase the atmospheric concentrations of both water and CO_2, producing the possibility of a runaway process that would be catastrophic from a human perspective. Fossil fuel combustion and deforestation during the last two hundred years have increased the atmospheric CO_2 concentration by 25%, and this increase is continuing. The same combustion processes responsible for the increasing atmospheric CO_2 concentration also introduce considerable quantities of particulate materials into the upper atmosphere. The effect of these would be to scatter more of the incoming solar radiation, reducing the amount that reaches and heats the earth's surface. The extent to which this process counteracts the greenhouse effect is still a matter of controversy; all that is known for sure is that the average temperature of the Earth is increasing.

Nitrous oxide

Carbon dioxide is not the only atmospheric gas of anthropogenic origin that can affect the heat balance of the earth; other examples are SO_2 and N_2O. Nitrous oxide is of particular interest, since its abundance is fairly high, and is increasing at a rate of about 0.5% per year. It is produced mainly by bacteria, and much of the increase is probably connected with introduction of increased nitrate into the environment through agricultural fertilization and sewage disposal. Besides being a strong infrared absorber, N_2O is

photochemically active, and can react with ozone. Any significant depletion of the ozone content of the upper atmosphere would permit more ultraviolet radiation to reach the earth. This would have numerous deleterious effects on present life forms, as well as contributing to a temperature increase. The warming effect attributed to anthropogenic additions of greenhouse gases to the atmosphere is estimated to be about 2 watts per m^2 , or about 1.5% of the 150 watts per m^2trapped by clouds and atmospheric gases. This is a relatively large perturbation compared to the maximum variation in solar output of 0.5 watts per m^2 that has been observed during the past century. Continuation of greenhouse gas emission at present levels for another century could increase the atmospheric warming effect by 6-8 watts per m^2.

> *A less-appreciated side effect of the increase in atmospheric carbon dioxide (and of other plant nutrients such as nitrates) may be reduction in plant species diversity by selectively encouraging the growth of species which are ordinarly held in check by other species that are able to grow well with fewer nutrients. This effect, for which there is already some evidence, could be especially pronounced when the competing species utilize the C_3 and C_4 photosynthetic pathways that differ in their sensitivity to CO_2.*

Contributors and Attributions

Stephen Lower, Professor Emeritus (Simon Fraser U.) Chem1 Virtual Textbook

This page titled 2.1.3.1: Structure and Composition of the Atmosphere is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.3.2: Origin and Evolution of the Atmosphere

Prebiotic atmosphere

The atmosphere of the Earth (and also of Venus and Mars) is generally believed to have its origin in relatively volatile compounds that were incorporated into the solids from which these planets accreted. Such compounds could include nitrides (a source of N_2), water (which can be taken up by silica, for example), carbides, and hydrogen compounds of nitrogen and carbon. Many of these compounds (and also some noble gases) can form clathrate complexes with water and some minerals which are fairly stable at low temperatures. The high temperatures developed during the later stages of accretion as well as subsequent heating produced by decay of radioactive elements presumably released these gases the surface. Even at the present time, large amounts of CO_2 , water vapor, N_2, HCl, SO_2 and H_2S are emitted from volcanos. The more reactive of these gases would be selectively removed from the atmosphere by reaction with surface rocks or dissolution in the ocean, leaving an atmosphere enriched in its present major components with the exception of oxygen which is discussed in the next section.

Any hydrogen present would tend to escape into space, causing the atmosphere to gradually became less reducing. However there is now some doubt that hydrogen and other volatiles (mainly the inert gases) were present in the newly accreted planets in anything like their cosmic abundance. The main evidence against this is the observation that gases such as helium, neon and argon, which are among the ten most abundant elements in the universe, are depleted on the earth by factors of 10^{-7} to 10^{-11} . This implies that there was a selective removal of these volatiles prior to or during the planetary accretion process.

The overall oxidation state of the earth's mantle is not consistent with what one would expect from equilibration with highly reduced volatiles, and there is no evidence to suggest that the composition of the mantle has not remained the same. If this is correct, then the primitive atmosphere may well have had about the same composition as the gases emitted by volcanos at the present time. These consist mainly of water and CO_2, together with small amounts of N_2, H_2, H_2S, SO_2, CO, CH_4, NH_3, HCl and HF. If water vapor was a major component of outgassing of the accreted earth, it must have condensed quite rapidly into rain. Any significant concentration of water vapor in the atmosphere would have led to a runaway greenhouse effect, resulting in temperatures as high as 400°C.

Origin of Atmospheric Oxygen

Free oxygen is never more than a trace component of most planetary atmospheres. Thermodynamically, oxygen is much happier when combined with other elements as oxides; the pressure of O_2 in equilibrium with basaltic magmas is only about 10^{-7} atm. Photochemical decomposition of gaseous oxides in the upper atmosphere is the major source of O_2 on most planets. On Venus, for example, CO_2 is broken down into CO and O_2 . On the earth, the major inorganic source of O_2 is the photolysis of water vapor; most of the resulting hydrogen escapes into space, allowing the O2 concentration to build up. An estimated 2 $\times$ 10^{11} g of O_2 per year is generated in this way. Integrated over the earth's history, this amounts to less than 3% of the present oxygen abundance. The partial pressure of O_2 in the prebiotic atmosphere is estimated to be no more than 10^{-3} atm, and may have been several orders of magnitude less. The major source of atmospheric oxygen on the earth is photosynthesis carried out by green plants and certain bacteria:

$$H_2O + CO_2 \longrightarrow (CH_2O)_x + O_2 \qquad (2.1.3.2.1)$$

A historical view of the buildup of atmospheric oxygen concentration since the beginning of the sedimentary record (3.7 $\times$ 10^9 ybp) can be worked out by making use of the fact that the carbon in the product of the above reaction has a slightly lower C^{13} content than does carbon of inorganic origin. Isotopic analysis of carbon-containing sediments thus provides a measure of the amounts of photosynthetic O_2 produced at various times in the past.

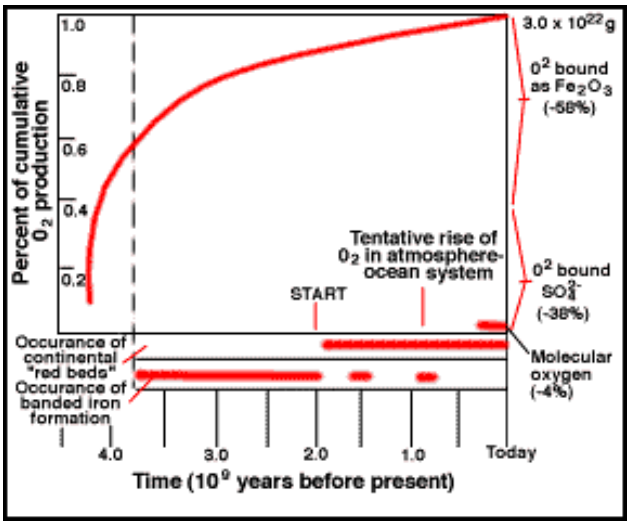


Figure 2.1.3.2.1: ***Evolution of atmospheric oxygen content.*** *Note carefully that the curve plots cumulative O_2 production, but that until a few hundred million years ago, most of this was taken up by Fe(II) compounds in the crust and by reduced sulfur; only after this massive "oxygen sink" became filled did free O_2 begin to accumulate in the atmosphere.*

Carbon Dioxide

Carbon dioxide has probably always been present in the atmosphere in the relatively small absolute amounts now observed (around 54 x 10^{15} mol = 54 Pmol). The reaction of CO_2 with silicate-containing rocks to form precambrian limestones suggest a possible moderating influence on its atmospheric concentration throughout geological time.

$$CaSiO_3 + CO_2 \rightarrow CaCO_3 + SiO_2 \tag{2.1.3.2.2}$$

About ten percent of the atmospheric CO_2 is taken up each year by photosynthesis. Of this, all except 0.05 percent is returned by *respiration*, almost entirely due to microorganisms. The remainder leaks into the slow part of the geochemical cycle, mostly as buried carbonate sediments.

Since the advent of the industrial revolution in 1860, the amount of CO_2 in the atmosphere has been increasing. Isotopic analysis shows that most of this has been due to fossil-fuel combustion; in recent years, the mass of carbon released to the atmosphere by this means has been more than ten times the estimated rate of natural removal into sediments. The large-scale destruction of tropical forests, which has accelerated greatly in recent years, is believed to exacerbate this effect by removing a temporary sink for CO_2.

The oceans have a large absorptive capacity for CO_2 owing to its reaction with carbonate:

$$CO_2 + {CO_3}^{2-} \rightleftharpoons 2\,{HCO_3}^{-} \tag{2.1.3.2.3}$$

There is about 60 times as much inorganic carbon in the oceans as in the atmosphere. However, efficient transfer takes place only into the topmost (100 m) wind-mixed layer of the ocean, which contains only about one atmosphere equivalent of CO_2. Further uptake is limited by the very slow transport of water into the deep ocean, which takes around 1000 years. For this reason, the buffering effect of the oceans on atmospheric CO_2 is not very effective; only about ten percent of the added CO_2 is taken up by the oceans.

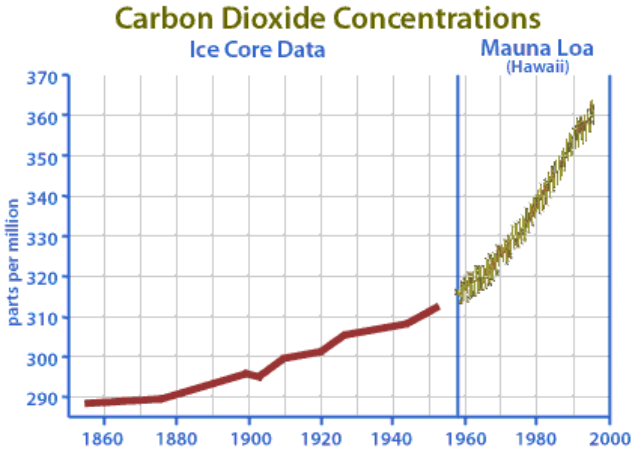


Figure. This set of plots, taken from an ETE page with data from NASA, shows the dramatic increase in atmospheric CO_2since the beginning of the industrial age. The "squiggles" on the Mauna Loa data reflect seasonal variations; more $CO_2$is taken up during the longer-daylight periods of the summer.

Contributors and Attributions

- Stephen Lower, Professor Emeritus (Simon Fraser U.) Chem1 Virtual Textbook

---

This page titled 2.1.3.2: Origin and Evolution of the Atmosphere is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

SECTION OVERVIEW

2.1.4: The Biosphere

The biosphere comprises the various regions near the earth's surface that contain and are dynamically affected by the metabolic activity of the approximately 1.5 million animal species and 0.5 million plant species that are presently known and are still being discovered at a rate of about 10,000 per year. The biosphere is the youngest of the dynamical systems of the earth, having had its genesis about 2 billion years ago. It is also the one that has most profoundly affected the other major environmental systems, particularly the atmosphere and the hydrosphere.

2.1.4.1: Chemistry and Energetics of the Life Process

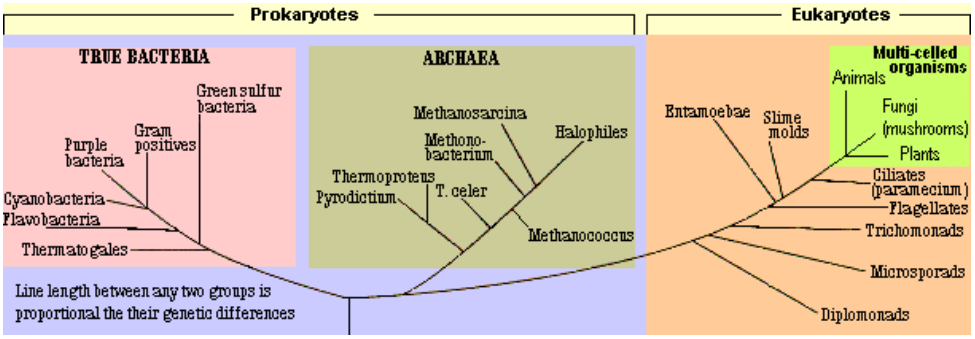
2.1.4.2: Biogeochemical Evolution

2.1.4.3: Gaia - Bioregulation of the Environment

This page titled 2.1.4: The Biosphere is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.4.1: Chemistry and Energetics of the Life Process

The biosphere comprises the various regions near the earth's surface that contain and are dynamically affected by the metabolic activity of the approximately 1.5 million animal species and 0.5 million plant species that are presently known and are still being discovered at a rate of about 10,000 per year. The biosphere is the youngest of the dynamical systems of the earth, having had its genesis about 2 billion years ago. It is also the one that has most profoundly affected the other major environmental systems, particularly the atmosphere and the hydrosphere.



"The tree of life"

About a third of the chemical elements cycle through living organisms, which are responsible for massive deposits of silicon, iron, manganese, sulfur, and carbon. Large quantities of methane and nitrous oxide are introduced into the atmosphere by bacterial action, and plants alone inject about 400,000 tons of volatile substances (including some metals) into the atmosphere annually.

> It has been suggested that biological activity might be responsible for the deficiency of hydrogen on Earth, compared to its very high relative abundance in the solar system as a whole. Bacteria capable of reducing hydrogen compounds into H_2 transform this element into a form in which it can escape from the earth; such bacteria may have been especially active in the reducing atmosphere of the early planet.
>
> A second mechamism might be the microbial production of methane, which presently injects about 10^9 T of CH_4 into the atmosphere each year. Some of this reaches the stratosphere, where it is oxidized to CO_2 and H_2O. The water vapor is photolyzed to H_2, which escapes into space. This may be the major mechanism by which water vapor (and thus hydrogen) is transported to the upper atmosphere; the low temperature of the upper atmosphere causes most of the water originating at lower levels to condense before it can migrate to the top of the atmosphere.

The increase in the abundance of atmospheric oxygen from its initial value of essentially zero has without question been the most important single effect of life on earth, and the time scale of this increase parallels the development of life forms from their most primitive stages up to the appearance of the first land animals about 0.5 billion years ago.

Bioenergetics

All life processes involve the uptake and storage of energy, and its subsequent orderly release in small steps during the metabolic process. This energy is taken up in the combination of ADP with inorganic phosphate to form ATP, in which form the energy is stored and eventually delivered to sites where it can provide the free energy needed for driving non-spontaneous reactions such as protein and carbohydrate synthesis.

$ADP + PO_4^{3-} \rightarrow ATP\ \Delta G° = +30\ kJ$

The three main metabolic processes are glycolysis, respiration, and photosynthesis. The first two of these extract free energy from glucose by breaking it up into smaller, more thermodynamically stable fragments. Photosynthesis reverses this process by capturing the energy of sunlight into ATP which then drives the buildup of glucose from CO_2.

Glycolysis and fermentation.

As its name implies, this most primitive (and least efficient) of all metabolic processes is based on the breakdown of a sugar into fragments having a smaller total free energy. Thus the 6-carbon sugar glucose can be broken down into two 3-carbon lactic acid units, or into three 2-carbon ethanol units.

$C_6H_{12}O_6 \rightarrow 2\ CH_3CHOHCOOH\ \Delta G° = –197\ kJ$

In this process, two molecules of ATP are produced, thus capturing 61 kJ of free energy. Since the standard free energy of glucose (with respect to its elements) is –2870 kJ, this represents an overall efficiency of about 2 percent.

The net reaction of glycolysis is essentially a rearrangement of the atoms initially present in the energy source. This is a form of fermentation, which is defined as the enzymatic breakdown of organic molecules in which other organic compounds serve as electron acceptors. Since there is no need for an external oxidizing or reducing agent, there is no change in the oxidation state of the environment.

Respiration

When the enzymatic degradation of organic molecules is accompanied by transfer of electrons to an external (and usually better) electron acceptor, the process is known as respiration. The overall reaction of respiration is the oxidation of glucose to carbon dioxide:

$C_6H_{12}O_6 + 6\ O_2 \rightarrow 6\ CO_2 + 6\ H_2O\ \Delta G° = –2380\ kJ$

In this process, 36 molecules of ADP are converted into ATP, thus capturing 1100 kJ of free energy: an efficiency of 38 percent.

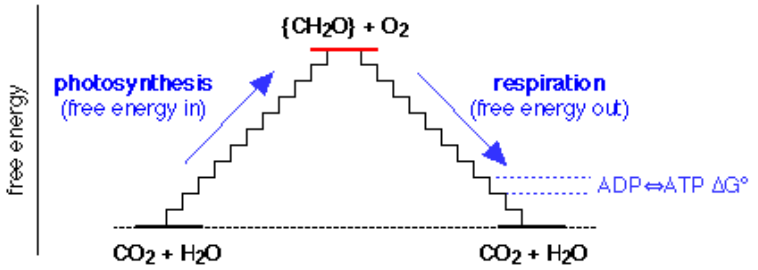
The oxidizing agent (electron sink) need not be oxygen; some bacteria reduce nitrates to NO or to N_2, and sulfates or sulfur to H_2S. These metabolic products can have far-reaching localized environmental effects, particularly if hydrogen ions are involved. Falling down the respiratory ladder presents a succinct picture of oxidation-reduction and the role of non-O_2 electron sinks in biological energy capture.

Photosynthesis

The energy of sunlight is trapped in the form of an intermediate which is able to deliver electrons to successively lower free-energy levels through the mediation of various molecules (mainly cytochromes) comprising an electron-transport chain. The free energy thus gained is utilized in part to reduce CO_2 to glucose, which is then available to supply metabolic energy by glycolysis or respiration. In green plants and eukaryotic algae, the source of hydrogen is water, the net reaction being

$6\ CO_2 + 6\ H_2O \rightarrow C_6H_{12}O_6 + 6\ O_2\ \Delta G° = +2830\ kJ$

For every CO_2 molecule fixed in this way, 469 kJ of free energy must be supplied. Red light of 680 nm wavelength has an energy of 168 kJ/mol; this implies that about three photons must be absorbed for every carbon atom taken up, but experiment indicates that about ten seem to be required.



The photosynthetic - respiration cycle.

Photosynthesis utilizes the energy of red light to add hydrogen (from H_2O) and electrons (from the O in H_2O) to CO_2, reducing it to **carbohydrate**. **Respiration** is the reverse of this process; electrons are removed from the carbohydrate (food is "burned") in small steps, each one releasing a small amount of energy. Some of this energy is liberated as heat, but part of it is used to add a phosphate group to ADP, converting it into ATP. **ATP** (adenosine triphosphate) is to an organism's energy needs as money is to our

material needs; it circulates to wherever it is required in order to bring about energy-requiring reactions or to make muscle cells contract. Each increment of energy given up by ATP converts it back to ADP and phosphate, ready to repeat the cycle.

Green plants are able to operate in both modes during the daylight hours, reverting to respiration-only at night. **Animals** carry out only the right side of the cycle and thus require as source of carbohydrate ("**food**") either directly (by eating plants), or indirectly by eating other animals that eat plants.

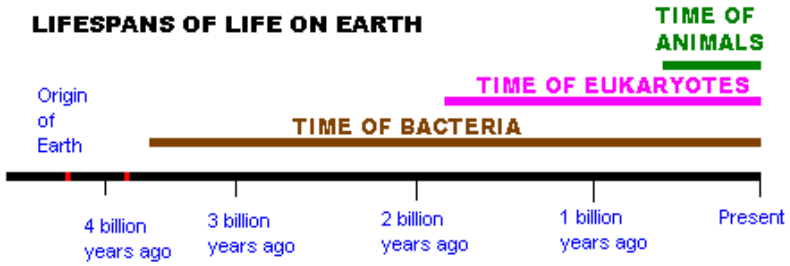
There are many kinds of photosynthetic bacteria, but with one exception (cyanobacteria) they are incapable of using water as a source of hydrogen for reducing carbon dioxide. Instead, they consume hydrogen sulfide or other reduced sulfur compounds, organic molecules, or elemental hydrogen itself, excreting the reducing agent in an oxidized state. Green plants, cyanobacteria, green filamentous bacteria and the purple nonsulfur bacteria utilize glucose by respiration during periods of darkness, while the green sulfur bacteria and the purple sulfur bacteria are strictly anaerobic.

This page titled 2.1.4.1: Chemistry and Energetics of the Life Process is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.4.2: Biogeochemical Evolution

Present evidence suggests that blue-green algae, and possibly other primitive microbial forms of life, were flourishing 3 billion years ago. This brackets the origin of life to within one billion years; prior to 4 billion years ago, surface temperatures were probably above the melting point of iron, and there was no atmosphere nor hydrosphere.

By about 3.8 billion years ago, or one billion years after the earth was formed, cooling had occurred to the point where rain was possible, and primitive warm, shallow oceans had formed. The atmosphere was anoxic and highly reducing, containing mainly CO_2, N_2, CO, H_2O, H_2S, traces of H_2, NH_3, CH_4, and less than 1% of the present amount of O_2, probably originating from the photolysis of water vapor. This oxygen would have been taken up quite rapidly by the many abundant oxidizable substances such as Fe(II), H_2S, and the like.



Timeline for development of the major life forms.

Evidence for early life

The fossil record that preserves the structural elements of organisms in sedimentary deposits has for some time provided a reasonably clear picture of the evolution of life during the past 750,000 years. In more recent years, this record has been considerably extended, as improved techniques have made it possible to study the impressions made by single-celled microorganisms embedded in rock formations.

The main difficulty in studying fossil microorganisms extending back beyond a billion years is in establishing that the relatively simple structural forms one observes are truly biogenic. There are three major kinds of evidence for this.

- Many of the most **primitive life forms are still thriving**, and these provide useful models with which some of the fossil forms can be compared.
- **Carbon isotope ratios** provide a second independent line of evidence for early life, or at least that of photosynthetic origin. In photosynthesis, $C^{12}O_2$ is taken up slightly more readily than is the heavier (and rarer) isotope $C^{13}O_2$; thus all but the very earliest life forms have left an isotopic fossil record even though the structural fossil may no longer be identifiable.
- A third evidence for early life is any indication of the **presence of free oxygen** in the local environment. Easily oxidizable species such as Fe(II) were very widely distributed on the primitive earth, and could not remain in contact with oxygen for very long without being oxidized. The oldest known formations of oxidized iron pyrite and of uranite are in sediments that were laid down between 2.0 and 2.3 billion years ago.

If all three of these lines of evidence are present in samples that can be shown to be contemporaneous with the sediments in which they are found, then the argument for life is incontrovertible. One of the most famous of these sites was discovered near Thunder Bay, Ontario in the early 1950's. The Gunflint Formation consists of an exposed layer of chert (largely silica) from which the overlying shale of the Canadian Shield had been removed. Microscopic examination of thin sections of this rock revealed a variety of microbial cell forms, including some resembling present freshwater blue-green algae. Also present in the Gunflint Deposits are the oldest known examples of metazoa, or organisms which display a clear differentiation into two or more types of cell. These deposits have been dated at 1.9-2.0 billion years.



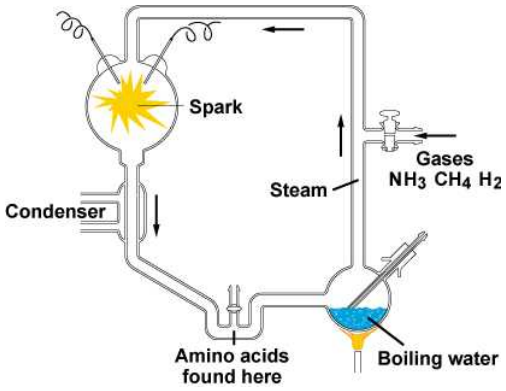
These filaments are believed to be the fossilized imprints of blue-green algae, one of the earliest life forms. They occur in the Bitter Springs Formation in Australia and are about 850 million years old.

The evidence from very old paleomicrobiotic deposits is less clear. Western Australia has yielded fossil forms that are apparently 2.8 billion years old, and other deposits in the same region contain structures resembling living blue-green algae. Other forms, heavily modified by chemical infiltration, bear some resemblence to a present iron bacterium, and are found in sediments laid down 3.5 billion years ago, but evidence that these fossils are contemporaneous with the sediments in which they are found is not convincing.

The oldest evidence of early life is the observed depletion of C^{13} in 3.8-billion year old rocks found in southwestern Greenland.

Origin of life: the organic environment

Under the conditions that prevailed at this time, most organic molecules would be thermodynamically stable, and there is every indication that a rich variety of complex molecules would be present. The most direct evidence of this comes from laboratory experiments that attempt to simulate the conditions of the primitive environment of this period, the first and most famous of these being the one carried out by Stanley Miller in 1953.



Schematic of the Miller experiment. *A mixture of the various reduced gases believed to be compose the early atmosphere circulate through an apparatus in which spark discharges (intended to simulate lightning) create a complex mixture of organic compounds. S.L. Miller: A production of amino acids under possible primitive earth conditions. 1953: Science (117) 528-529. U. Texas provided the above diagram, and contains a very good description of of primitive life with many pictures of fossil organisms.*

Since that time, other experiments of a similar nature have demonstrated the production of a wide variety of compounds under prebiotic conditions, including nearly all of the monomeric components of the macromolecules present in living organisms. In addition, small macromolecules, including peptides and sugars, as well as structural entities such as lipid-based micelles, have been prepared in this way.

The discovery in 1989 of a number of amino acids in the iridum-rich clay layer at the Cretaceous-Tertiary boundary suggests that bio-precursor molecules can be formed or deposited during a meteoric impact. Although this particular event occurred only 65 million years ago (and is presumed to be responsible for the extinction of the dinosaurs), the Earth has always been subject to meteoric impacts, and it is conceivable that these have played a role in the origin of life.

The presence of clays, whose surfaces are both asymmetric and chemically active, could have favored the formation of species of a particular chirality; a number of experiments have shown that clay surfaces can selectively adsorb amino acids which then form small peptides. It has been suggested that the highly active and ordered surfaces of clays not only played a crucial role in the formation of life, but might have actually served as parts of the first primitive self-replicating life forms, which only later evolved into organic species.

The first organisms

Since no laboratory experiment has yet succeeded in producing a self-replicating species that can be considered living, the mechanism by which this came about in nature must remain speculative. Infectious viruses have been made in the laboratory by simply mixing a variety of nucleotide precursors with a template nucleic acid and a replicase enzyme; the key to the creation of life is how to do the same thing without the template and the enzyme.

Smaller polynucleotides may have formed adventitiously, possibly on the active surface of an inorganic solid. These could form complementary base-paired polymers, which might then serve as the templates for larger molecules. Non-enzymatic template-directed synthesis of nucleotides has been demonstrated in the laboratory, but the resulting polymers have linkages that are not present in natural nucleotides.

It has been suggested that these linkages could have been selectively hydrolyzed by a long period of cycling between warm, cool, wet, and dry environmental conditions. The earth at that time was rotating more rapidly than it is now; cycles of hydration-dehydration and of heating-cooling would have been more frequent and more extreme.

The first organisms would of necessity have been heterotrophs— that is, they derived their metabolic energy from organic compounds in the environment. Their capacity to synthesize molecules was probably very limited, and they would have had to absorb many key substances from their surroundings in order to maintain their metabolic activity. Among the most primitive organisms of this kind are the archaeons, which are believed to be predecessors of both bacteria and eucaryotes. DNA sequencing of one such organism, a methane-producer that lives in ocean-bottom sediments at 200 atm and 48-94°C, reveals that only about a third of the genes resemble those of bacteria or eucaryotes.

It has been estimated that about 50 genes are required in order to define the minimal biochemical and structural machinery that a hypothetical simplest possible cell would have.

Development of photosynthetic organisms

The earliest organisms derived their metabolic energy from the organic substances present in their environment; once they began to reproduce, this nutrient source began to become depleted.

Some species had probably by this time developed the ability to reduce carbon dioxide to methane; the hydrogen source could at first have been H_2 itself (at that time much more abundant in the atmosphere), and later, various organic metabolites from other species could have served.

Before the food supply neared exhaustion, some of these organisms must have developed at least a rudimentary means of absorbing sunlight and using this energy to synthesize metabolites. The source of hydrogen for the reduction of CO_2 was at first small organic molecules; later photosynthetic organisms were able to break this dependence on organic nutrients and obtain the hydrogen from H_2S.

These bacterial forms were likely the dominant form of life for several hundred million years. Eventually, due perhaps to the failing supply of H_2S, plants capable of mediating the photochemical extraction of hydrogen from water developed. This represented a large step in biochemical complexity; it takes 10 times as much energy to abstract hydrogen from water than from hydrogen sulfide, but the supply is virtually limitless.

It appears that photosynthesis evolved in a kind of organism whose present-day descendents are known as cyanobacteria.

Procaryotes and eucaryotes

The five “kingdoms” into which living organisms are classified are Monera, Protista (protozoans, algae), Fungi, Plantae, and Animalia. The genetic (and thus, evolutionary) relations between these and the subcategories within them are depicted below.

Superimposed on this, however, is an even more fundamental division between the procaryotes and eucaryotes.

Procaryotes

In this group are primitive organisms whose single cells contain no nucleus; the gene-bearing structure is a single long DNA chain that is folded irregularly throughout the cell. Procaryotic cells usually reproduce by budding or division; where sexual reproduction does occur, there is a net transfer of some genetic material from one cell to another, but there is never an equal contribution from both parents.

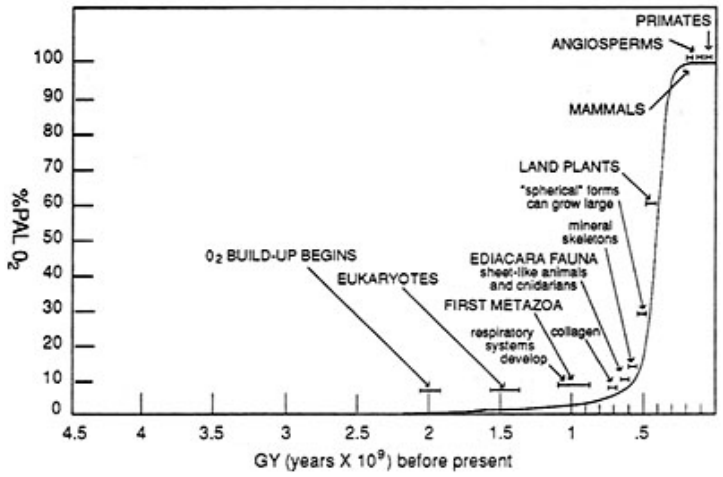
In spite of their primitive nature, procaryotes constitute the majority of organisms in the biosphere. The division between bacteria and archaea within the procaryotic group is a fairly recent one. Archaea are now believed to be the most primitive of all organisms, and include the so-called extremophiles that occupy environmental niches in which life was at one time thought to be impossible; they have been found in sedimentary rocks, hot springs, and highly saline environments.

Eucaryotes.

All other organisms— seaweeds (algae), protozoa, molds, fungi, animals and plants, are composed of eucaryotic cells. These all have a membrane-bound nucleus, and with a few exceptions they all reproduce by mitosis, in which the chromosomes split longitudinally and move toward opposite poles. Other organelles unique to eucaryotes are mitochondria, ribosomes, and structural elements such as microtubules.

Oxygen and biogeochemical evolution

Oxygen is poisonous to all forms of life in the absence of enzymes that can reduce the highly reactive byproducts of oxidation and oxidative metabolism (peroxides, superoxides, etc.). All organic compounds are thermodynamically unstable in the presence of oxygen; carbon-carbon double bonds in lipids are subject to rapid attack. Prebiotic chemical evolution leading to the development of biopolymers was possible only under the reducing, anoxic conditions of the primitive atmosphere.



The rise of atmospheric oxygen. *Once organisms existed that could use water as a hydrogen source for the reduction of carbon dioxide, O_2began to be introduced into the atmosphere. The widespread occurrence of ferrous compounds in surface rocks and sediments provided a sink for this oxygen that probably did not become saturated until about 2 billion years ago, when the atmospheric oxygen abundance first rose above about 1 percent.Illustration from All about entropy, the laws of thermdynamics, and order from disorder. There is lots of good stuff here!*

As the oxygen concentration began to rise, organisms in contact with the atmosphere had to develop protective mechanisms in order to survive. One indication of such adaptation is the discovery of fossil microbes whose cell walls are unusually thick. A more useful kind of adaptation was the synthesis of compounds that would detoxify oxygen by reacting rapidly either with O_2 itself or with peroxides and other active species derived from it. Isoprenoids (the precursors of steroids) and porphyrins are examples of two general classes of compounds that are found in nearly all organisms, and which may have originated in this way. Later, highly efficient oxygen mediating enzymes such as peroxidase and catalase developed. The widespread phenomenon of bioluminescence may be the result of a very early adaptation to oxygen. The compound luciferin is a highly efficient oxygen detoxifier, which also

happens to be able to emit light under certain conditions. Bioluminescence probably developed as a by-product in early procaryotic organisms, but was gradually lost as more efficient detoxifying mechanisms became available.

Development of the eucaryotes

In spite of the deleterious effects of oxygen on cell biomolecules, O_2 is nevertheless an excellent electron sink, capable of releasing large quantities of energy through the oxidation of glucose. This energy can be efficiently captured through oxidative phosphorylation, the key process in respiration.

A cell that utilizes oxygen must have a structural organization that isolates the oxygen-consuming respiratory centers from the other parts of the cell that would be poisoned by oxygen or its reaction products. Some procaryotic organisms have developed in this way; a number of cyanobacteria and other species are facultative anaerobes which can survive both in the presence and absence of oxygen.

It is in the eucaryotic cell, however, that this organization is fully elaborated; here, respiration occurs in membrane-bound organelles called mitochondria. With only a few exceptions, all eucaryotic organisms are obligate aerobes; they can rarely survive and can never reproduce in the absence of oxygen. Mitotic cell division depends on the contractile properties of the protein actomyosin, which only forms when oxygen is present.

The development of the eucaryotic cell about 1.4 billion years ago is regarded as the most significant event in the evolution of the earth and of the biosphere since the appearance of photosynthesis and the origin of life itself. How did it come about? The present belief, supported by an increasing amount of evidence, suggests that it began when one species of organism engulfed another. The ingested organism possessed biochemical machinery not present in the host, but which was retained in such a way that it conferred a selective evolutionary advantage on the host. Eventually the two organisms became able to reproduce as one, and so effectively became a single organism. This process is known as *endosymbiosis*

According to this view, mitochondria represent the remains of a primitive oxygen-tolerant organism that was incorporated into one that could produce the glucose fuel for the oxygen to burn. Chloroplasts were once free-living photosynthesizing procaryotes similar to present-day cyanobacteria. It is assumed that some of these began parasitising respiratory organisms, conferring upon them the ability to synthesize their carbohydrate food during daylight. The immense selective advantage of this arrangement is evident in the extent of the plant kingdom.

Consequences of the oxygen increase

It is interesting that an atmospheric oxygen concentration of about 1 percent, known as the *Pasteur point* is both the maximum that obligate anaerobes can tolerate, and the minimum required for oxidative phosphorylation.

> *Louis Pasteur discovered that some bacteria are anaerobic and unable to tolerate oxygen above 1% concentration.*

As was mentioned previously, the oxygen produced by the first photosynthetic organisms was taken up by ferrous iron in sediments and surface minerals. The widespread deposits known as banded iron formations consist of alternating layers of Fe(III)-containing oxides (hematite and magnetite) that were laid down between 1 and 2 billion years ago; the layering may reflect changing climatic or other environmental conditions that brought about a cycling of the organism population.

During the buildup of oxygen, an equivalent amount of carbon had to be deposited in sediments in order to avoid the thermodynamically spontaneous back reaction which would consume the O_2through oxidation of the organic matter. Thus the present levels of atmospheric oxygen are due to a time lag in the geochemical cycling of photosynthetic products.

As the oxygen concentration increased, evolution seems to have speeded up; this may reflect both the increased metabolic efficiency and the greater biochemical complexity of the eucaryotic cell. The oldest metazoan (multiple-celled) fossils are coelenterates that appeared about 700 million years ago. Modern representatives of this group such as marine worms and jellyfish can tolerate oxygen concentrations as low as 7%, thus placing a lower boundary on the atmospheric oxygen content of that era. The oldest fossil organisms believed to have possessed gills, which function only above 10% oxygen concentration, appeared somewhat later.Carbon dioxide decreased as oxygen increased, as indicated by the prevalence of dolomite over limestone in early marine sediments.

This page titled 2.1.4.2: Biogeochemical Evolution is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.1.4.3: Gaia - Bioregulation of the Environment

The Gaia Hypothesis

The physical conditions under which life as we know it can exist encompass a relatively narrow range of temperature, pH, osmotic pressure, and ultraviolet radiation intensity. It seems remarkable enough that life was able to get started at all; it is even more remarkable that it has continued to thrive in the face of all the perils that have, or could have occurred, during the past 3 billion years or so.

During the time that life has been evolving, the sun has also been going through the process of evolution characteristic of a typical star; one consequence of this is an increase in its energy output by about 30 percent during this time. If the sun's output should suddenly drop to what it was 3 billion years ago, the oceans would freeze. How is it that the earth was not in a frozen state for the first 1.5 billion years of life's existence? Alternatively, if conditions were somehow suitable 3 billion years ago, why have the oceans not long since boiled away?

A rather non-traditional answer to this kind of question is that the biosphere is far from playing a passive role in which it is continually at the mercy of environmental conditions. Instead, the earth's atmosphere, and to a lesser extent the hydrosphere, may be actively maintained and regulated by the biosphere. This view has been championed by the British geochemist J.E. Lovelock, and is known as the *Gaia hypothesis*.

> *Gaia is another name for the Greek earth-goddess Ge, from which root the sciences of geography, geometry, and geology derive their names. Lovelock's book Gaia: a new look at life on Earth (Oxford, 1979) is a short and highly readable discussion of the hypothesis.*

Evidence in support of this hypothesis is entirely circumstantial, but nevertheless points to important questions that must be answered: how have the climatic and chemical conditions on the earth remained optimal for life during all this time; how can the chemical composition of the atmosphere remain in a state that is tens of orders of magnitude from equilibrium?

Note

Although the Gaia hypothesis has received considerable publicity in the popular press, it has never been very well received by the scientific community, many of whom feel that there is no justification for proposing a special hypothesis to describe a set of connections which can be quite adequately explained by conventional geochemical processes. More recently, even Lovelock has backed away from the teleological interpretation of these relations, so that the Gaia hypothesis should now be more properly described as a set of loosely connected effects, rather than as a hypothesis. Nevertheless, these effects and the mechanisms that might act to connect them are sufficiently interesting that it seems worthwhile to provide an overview of the major observations that led to the development of the hypothesis.

Teleology is the doctrine that natural processes operate with a purpose. See *No longer willful, Gaia becomes respectable*. 1988: Science 240 393-395.

Bioregulation of the Atmosphere

The increase in the oxygen content of the atmosphere as a result of the development of the eucaryotic cell was discussed above. Why has the oxygen content leveled off at 21 percent? It is interesting to note that if the oxygen concentration in the atmosphere were only four percent higher, even damp vegetation, once ignited by lightning, would continue to burn, enveloping vast areas of the earth in a firestorm. Evidence for such a worldwide firestorm that may be related to the extinction of the dinosaurs has recently been discovered. The charcoal layers found in widely distributed sediments laid down about 65 million years ago are coincident with the iridium anomaly believed to be due to the collision of a large meteor with the earth.

- **Oxygen**: Regulation of the oxygen partial pressure is probably achieved by a balance between its production through photosynthesis and its consumption during oxidation of organic matter; the present steady state requires the burial of about 0.1% the carbon that is fixed annually, leaving one O_2 molecule in the air for each atom of carbon removed from the photosynthetic cycle. The large quantities of microbially-produced methane and N_2O also constitute important oxygen sinks; if methanogenic bacteria should suddenly cease to exist, the O_2 concentration would rise by 1% in about 12,000 years. This type

of regulation implies a negative feedback mechanism, in which an increase in atmospheric oxygen would increase the activity of organisms capable of generating metabolic products that react with it.

- **Nitrous oxide**: Nitrous oxide, in addition to serving as an oxygen sink, might also be a factor in the regulation of the intensity of the ultraviolet component of sunlight. N_2O acts as a catalytic intermediate in the decomposition of stratospheric ozone, which shields the earth from excessive ultraviolet radiation.
- **Ammonia**: Ammonia, another atmospheric gas, is produced by the biosphere in approximately the same quantities as methane, 10^9 tons per year, and at the expense of a considerable amount of metabolic energy. The function of NH_3 could well be to regulate the pH of the environment; in the absence of ammonia, the large amounts of SO_2 and HCl produced by volcanic action would reduce the pH of rain to about 3. The fact that the atmospheric concentration of ammonia is only 10^{-8} times that of N_2 should not imply that this "trace" component plays a less significant role in the overall nitrogen cycle than does than N_2. In fact, the annual rates of production of the two gases are roughly the same; the much lower steady-state concentration of NH_3 is due to its faster turnover time.
- **Nitrogen**: As stable as the triply-bonded N_2 molecule is, there is a still more stable form of nitrogen: the hydrated nitrate ion. How is this stability consistent with the predominance of nitrogen in the atmosphere? The answer is that it is not: if it were not for nitrogen-fixing bacteria (powered directly or indirectly by the free energy of ATP captured from sunlight), the nitrogen content of the atmosphere would disappear to almost zero. This would raise the oxygen fraction to disastrously high levels, and the additional NO_3^- concentration would increase the ionic strength and osmotic pressure of seawater to levels inconsistent with most forms of life.

Bioregulation of the Oceans

The input of salts into the sea from streams and rivers is about 5.4×10^8 tons per year, into a total volume of about 1.2×10^9 km^3 yr^{-1} . Upwelling of juvenile water and hydrothermal action at oceanic ridges provide additional inputs of salts. With a few bizarre exceptions such as the brine shrimp and halophilic bacteria, 6 percent is about the maximum salinity level that organisms can tolerate. The internal salinities of cells must be maintained at much lower levels (around 1%) to prevent denaturation of proteins and other macromolecules whose conformations are dependent on electrostatic forces. At higher levels than this, the electrostatic interaction between the salt ions and the cell membrane destroys the integrity of the latter so that it can no longer pump out salt ions that leak in along the osmotic gradient. At the present rate of salt input, the oceans would have reached their present levels of salinity millions of years ago, and would by now have an ionic strength far to high to support life, as is presently the case in the landlocked Dead Sea.

> *The present average salinity of seawater is 3.4 percent. The salinity of blood, and of many other intra- and intercellular fluids in animals, is about 0.8 percent. If we assume that the first organisms were approximately in osmotic equilibrium with seawater, then our body fluids might represent "fossilized" seawater as it existed at the time our predecessors moved out of the sea and onto the land.*

By what processes is salt removed from the oceans in order to maintain a steady-state salinity? This remains one of the major open questions of chemical oceanography. There are a number of answers, mostly based on strictly inorganic processes, but none is adequately supported by available evidence. For example, Na^+ and Mg^{2+} ions could adsorb to particulate debris as it drops to the seafloor, and become incorporated into sediments. The requirement for charge conservation might be met by the involvement of negatively charged silicate and hydroxyaluminum ions. Another possible mechanism might be the burial of salt beds formed by evaporation in shallow, isolated arms of the sea, such as the Persian Gulf. Extensive underground salt deposits are certainly found on most continents, but it is difficult to see how this very slow mechanism could have led to an unfluctuating salinity over shorter periods of highly variable climatic conditions.

The possibility of biological control of oceanic salinity starts with the observation that about half of the earth's biomass resides in the sea, and that a significant fraction of this consists of diatoms and other organisms that build skeletons of silica. When these organism die, they sink to the bottom of the sea and add about 300 million tons of silica to sedimentary rocks annually. It is for this reason that the upper levels of the sea are undersaturated in silica, and that the ratio of silica to salt in dead salt lakes is much higher than in the ocean.

These facts could constitute a basis for a biological control of the silica content of seawater; any link between silica and salt could lead to the control of the latter substance as well. For example, the salt ions might adsorb onto the silica skeletons, and be carried down with them; if the growth of these silica-containing organisms is itself dependent on salinity, we would have our negative feedback mechanism.

The continual buildup of biogenic sedimentary deposits on the ocean floor might possibly deform the thin oceanic crust by its weight, and cause local heating by its insulating properties. This could conceivably lead to volcanic action and the formation of new land mass, thus linking the lithosphere into Gaia.

Contributors and Attributions

- Stephen Lower, Professor Emeritus (Simon Fraser U.) Chem1 Virtual Textbook

This page titled 2.1.4.3: Gaia - Bioregulation of the Environment is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.2: Supplemental Modules (Environmental Chemistry)

2.2: Supplemental Modules (Environmental Chemistry) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

SECTION OVERVIEW

2.2.1: Acid Rain

"Normal" rainfall is slightly acidic because of the presence of dissolved carbonic acid. Carbonic acid is the same as that found in soda pop.

2.2.1.1: Acid Rain

2.2.1.2: Acid Rain Transport

2.2.1.3: Acid Snow

2.2.1.4: Electricity Generation

2.2.1.5: Playground 1

2.2.1.6: Sources of Nitrogen Oxides

2.2.1.7: Sources of Sulfur Oxides

2.2.1: Acid Rain is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2.1.1: Acid Rain

"Normal" rainfall is slightly acidic because of the presence of dissolved carbonic acid. Carbonic acid is the same as that found in soda pop.

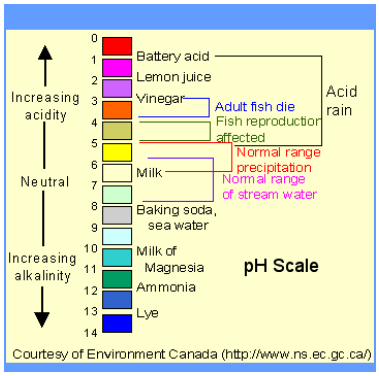
Introduction

The pH of "normal" rain has traditionally been given a value of 5.6. However scientists now believe that the pH of rain may vary from 5.6 to a low of 4.5 with the average value of 5.0. Acid rain or acid snow is a direct result of the method that the atmosphere cleans itself. The tiny droplets of water that make up clouds, continuously capture suspended solid particles and gases in the atmosphere. The gases of sulfur oxides and nitrogen oxides are chemically converted into sulfuric and nitric acids. The non-metal oxide gases react with water to produce acids (ammonia produces a base).

$$SO_2 + HOH \rightarrow H_2SO_3 \tag{2.2.1.1.1}$$

$$2NO_2 + HOH \rightarrow HNO_2 + HNO_3 \tag{2.2.1.1.2}$$

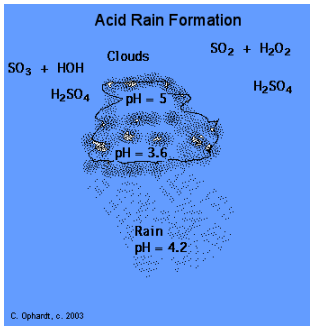
When enough of the tiny cloud droplets clump together to form a larger water drop it may fall to the earth as "wet" acid precipitation including rain, snow, ice, sleet, or fog. See Acids and Bases for more details.



Formation

The sulfuric and nitric acids formed from gaseous pollutants can easily make their way into the tiny cloud water droplets. These sulfuric acid droplets are one component of the summertime haze in the eastern United States. Some sulfuric acid is formed directly in the water droplets from the reaction of sulfur dioxide and hydrogen peroxide. Some of these sulfuric acid particles drop to the earth as "dry" acid deposition. The tiny water droplets containing sulfuric acid provide a ready surface to attract more molecules of water to form a larger droplet of dilute sulfuric acid.

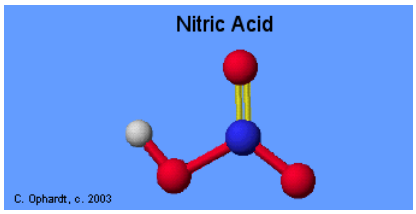
Water droplets collected from the base of clouds in the Eastern U.S.during the summer have an average pH of 3.6, with some values as low as pH 2.6. The pH in the upper portion of a cloud is much higher. The final rain droplet has an average pH of 4.2 in the Northeast U.S. In Los Angeles, the pH of fog has been measured at 2.0 - about the acidity of lemon juice.



Contrary View

The benchmark of "natural rain" is 5.6. Acid precipitation in the range of 4.2-5.0 has been recorded in most of the Eastern United States and Canada. EPRI (Electric Power Research Institute) likes to compare these values to familiar objects to give the impression that these pH values are not harmful. Examples: Carrots = 5.0, Bananas = 4.6, Tomatoes = 4.2, Apples and soft drinks = 3.0, Lemon juice = 2.0. EPRI also contends that pH 5.6 may or may not be a valid reference point. It should not be considered the "background" or "natural" acidity of precipitation.

Even without man-made influences, there are natural sources of sulfur oxides, nitrogen oxides, and other species important to determining the precipitation acidity at any given time. Hence, trying to quantify man's contribution to the natural condition will never be possible, since the "natural background" condition cannot be known.



In the forest areas of Brazil at the headlands of the Amazon River, an area remote from civilization, the monthly average of 100 rain events in the 1960s ranged from pH 4.3 to pH 5.2, with the median value of pH 4.6 and one reading as low as pH 3.6. On the island of Hawaii, remote from all industrial activity, the weighted average of precipitation over a 4 year period was pH 5.3, with a minimum value of pH 3.8.

Outside Links

- www.epa.gov/acidrain/what/index.html
- Bretherick, Leslie. "Nitric acid, nitrates, and nitro compounds: Energetic servants, but overwhelming masters (SAFETY)." *J. Chem. Educ.* **1989**, *66*, A220.
- Driscoll, Jerry A. "Acid Rain Demonstration: The Formation of Nitrogen Oxides as a By-Product of High-Temperature Flames in Connection with Internal Combustion Engines." *J. Chem. Educ.* **1997** *74* 1424.
- Johns, Nicholas; Longstaff, Stephen J. "A convenient, low-cost method for determining sulfate in acid rain (F&R)." *J. Chem. Educ.* **1987**, *64*, 449.

Contributors

- Charles Ophardt, Professor Emeritus, Elmhurst College
- Virtual Chembook

2.2.1.1: Acid Rain is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

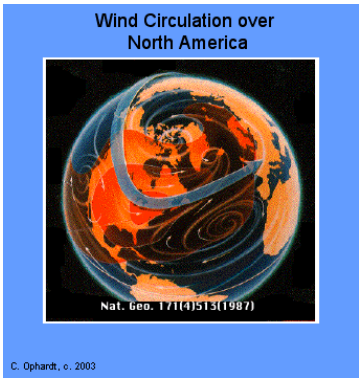
2.2.1.2: Acid Rain Transport

The reactions of sulfur oxides to form sulfuric acid are quite slow. Sulfur dioxide may remain airborne for 3-4 days.As a consequence acid rain derived from sulfur oxides may travel for hundreds of miles or even a thousand miles. Nitrogen oxides may persist for only about one half day and therefore may travel only tens or hundreds of miles.

Introduction

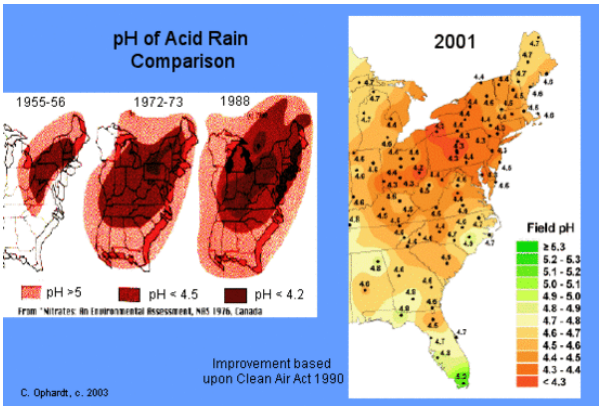
Once airborne, the sulfur and nitrogen oxides eventually come down in one form or another. Where they come down depends on the height of the smokestack and the prevailing weather conditions. In general, prevailing winds in North America transport pollutants from west to east or northeast. The nine largest coal burning states are in the Midwest and the Ohio River valley. It is estimated that two thirds of the acid rain in the Northeast and Eastern Canada comes from these sources.

Blue arrow shows the upper winds that travel from the west to the east or northeast. Winds travel from the mid-west to the northeast. In addition, a copper-nickel smelter in Sudbury, Ontario, just north of Lake Huron is the most significant sulfur oxide source in Canada. The winds may also carry the sulfur oxide clouds to the Northeast in the U.S. where it may be converted to acid rain.



pH of Acid Rain over a Period of 50 years

These maps show how the areas of lower pH have spread in a 30 years 1955-1988. Darkest area is lowest pH. Since the Clean Air Act Amendments of 1990, there have been significant decreases in the amount of sulfur oxides escaping from the electric power plants. As a result there has been a measurable reduction in the amount of acid rain, which is actually translated as an increase in the pH levels (higher pH means less acid).



Contributors

- Charles Ophardt, Professor Emeritus, Elmhurst College, Virtual Chembook

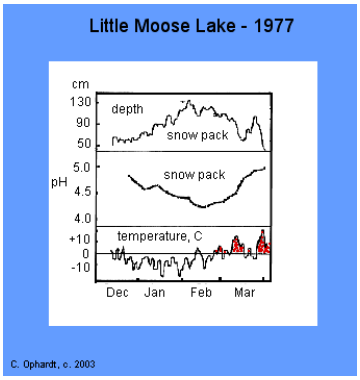
2.2.1.2: Acid Rain Transport is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2.1.3: Acid Snow

The impact of acid precipitation on aquatic ecosystems may be intensified by melting snow. When snow melts rapidly in the spring, the stream or lake may be "shocked" with an excessive amount of acid. In the spring, at the time of acid snow melting, the various aquatic organisms are reproducing and are the most sensitive increases in acid.

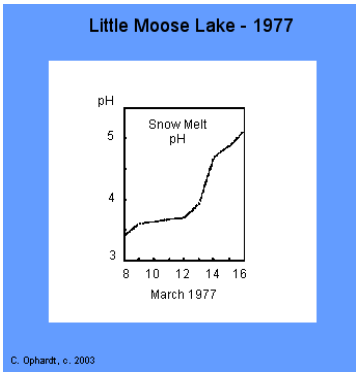
Little Moose Lake

During the winter of 1976-77, there was no significant snow melt from mid-December through February at Little Moose Lake, New York. The highest depth was 130 cm. (top graph). The snow had an average pH of 4.4 in Feb. (middle graph). The snow started to melt in early March.(bottom graph).



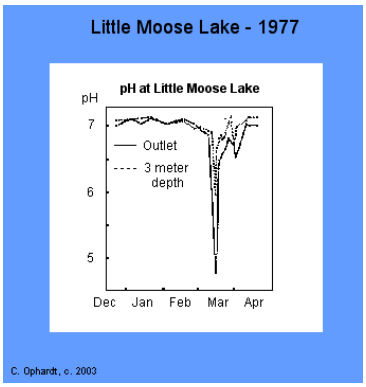
Snow Thaw at Little Moose Lake

The first major thaw in early March resulted in the release of 80% of the stored acid in a one week period. During the first couple of days, the snow melt water pH was 3.4-3.6. The pH gradually rose over the eight day period to pH 5.0.



pH at Little Moose Lake

The pH levels in Little Moose lake are normally about 7.0. During the snow melting, in early March, the lake pH dropped to 6.0. An outlet stream from the lake reached a low pH of 4.8. A small brook nearby hit a low pH of 4.6 during the snow melt period. The average pH in this brook during the rest of the year is about 5.4. There was a definite relationship of the amount of aluminum in the brook vs. pH. As the pH decreased, the aluminum concentration increased.



Captive populations of adult, year old, and larval brook trout had been maintained over the winter in Little Moose Lake water without any problems. During the snow melt in March, 3 adult brook trout died, 25 one year old trout died, and about 50,000 recently hatched trout died. Various other abnormal behaviors were also observed.

Contributors

- Charles Ophardt, Professor Emeritus, Elmhurst College, Virtual Chembook

---

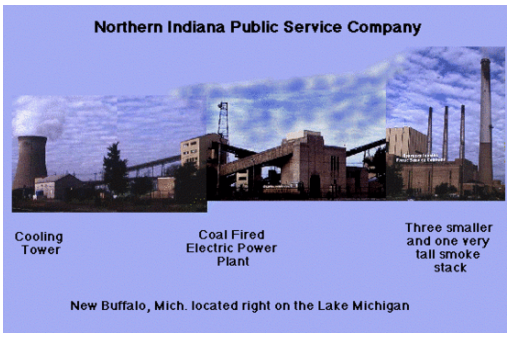
2.2.1.3: Acid Snow is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2.1.4: Electricity Generation

Electricity is produced at a an electric power plant. Some fuel source, such as coal, oil, natural gas, or nuclear energy produces heat, which is used to boil water to create steam. The steam under high pressure is then used to spin a turbine that interacts with a system of magnets to produce electricity. The electricity is transmitted as moving electrons through a series of wires to homes and business.

Introduction

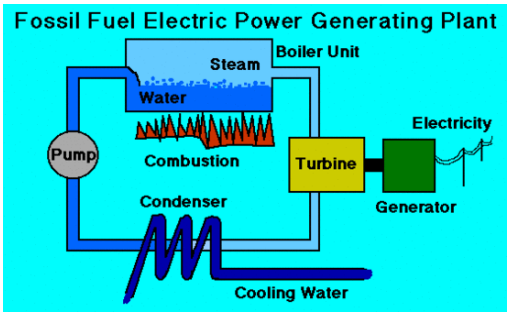
This is a typical electric power plant located in Shawville, Pennsylvania.



Notice the large pile of coal on the left side of the plant and the three smokestacks, each one taller than the previous. The tallest stack was built to cut down on the local air pollution, where the sulfur oxides are emitted higher into the atmosphere. This has not proven to be a solution to the probelm. As a result the sulfur oxides now travel great distances before coming down in the form of acid rain.

Electric Power Plants

Electric Power Plants have a number of components in common and are an interesting study in the various forms and changes of energy necessary to produce electricity.



- **Boiler Unit:** Almost all of power plants operate by heating water in a boiler unit into super heated steam at very high pressures. The source of heat from combustion reactions may vary in fossil fuel plants from the source of fuels such as coal, oil, or natural gas. Biomass or waste plant parts may also be used as a source of fuel. In some areas solid waste incinerators are also used as a source of heat. All of these sources of fuels result in varying amounts of air pollution, as well as, the carbon dioxide (a gas implicated in global warming problems).
- **Turbine-Generator**: The super heated steam is used to spin the blades of a turbine, which in turn is used in the generator to turn a coil of wires within a circular arrangements of magnets. The rotating coil of wire in the magnets results in the generation of electricity.
- **Cooling Water**: After the steam travels through the turbine, it must be cooled and condensed back into liquid water to start the cycle over again. Cooling water can be obtained from a nearby river or lake. The water is returned to the body of water 10-20 °C higher in temperature than the intake water. Alternate method is to use a very tall cooling tower, where the evaporation of water falling through the tower provides the cooling effect.
- In a **nuclear power plant**, the fission chain reaction of splitting nuclei provides the source of heat.

Creating Electricity using a Generator

If a magnetic field can create a current then we have a means of generating electricity. Experiments showed that a magnetic just sitting next to a wire produced no current flow through that wire. However, if the magnet is moving, a current is induced in the wire. The faster the magnet moves, the greater the induced current. This is the principal behind simple electric generators in which a wire loop is rotated between to stationary magnetics. This produces a continuously varying voltage which in turn produces an alternating current .

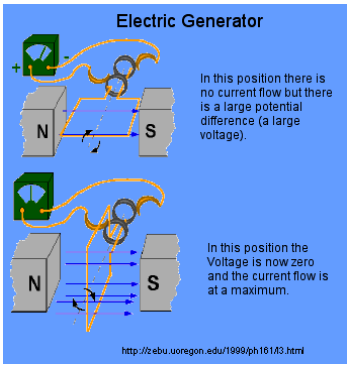


Diagram of a simple electric generator is shown above. To generate electricty then, some (mechanical) mechanism is used to turn a crank that rotates a loop of wire between stationary magnets. The faster the crank turns, the more current that is generated. In hydroelectric, the falling water turns the turbine. The wind can also turn the turbine. In fossil fuel plants and nuclear plants, water is heated to steam which turns the turbine.

Outside Links

- http://zebu.uoregon.edu/1999/ph161/l3.html

Contributors

- Charles Ophardt, Professor Emeritus, Elmhurst College, Virtual Chembook

2.2.1.4: Electricity Generation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2.1.5: Playground 1

2.2.1.5: Playground 1 is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.2.1.6: Sources of Nitrogen Oxides

Natural Sources - Lighting Bolts

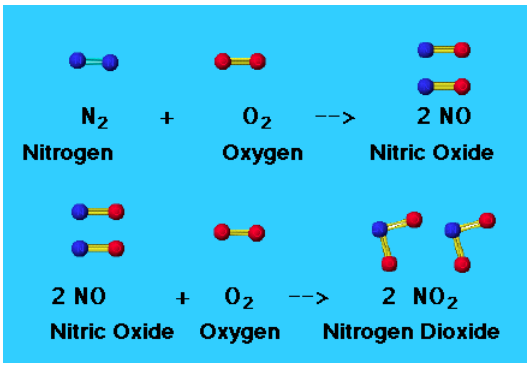
A natural source of nitrogen oxides occurs from a lightning stroke. The very high temperature in the vicinity of a lightning bolt causes the gases oxygen and nitrogen in the air to react to form **nitric oxide.**

$$N_2 + O_2 \longrightarrow NO \tag{2.2.1.6.1}$$

The nitric oxide very quickly reacts with more oxygen to form **nitrogen dioxide.**

$$NO + O_2 \longrightarrow NO_2 \tag{2.2.1.6.2}$$

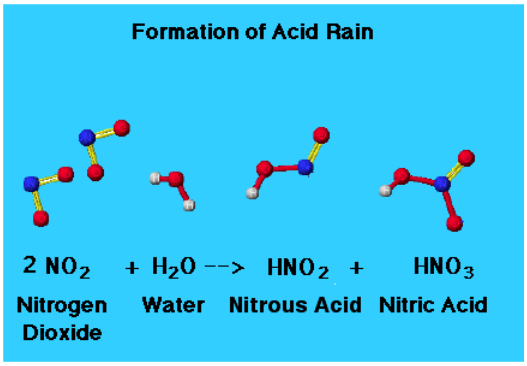
Both of the nitrogen compounds are known collectively as nitrogen oxides or NO_x.



Human Sources of Nitrogen Oxides

At normal temperatures the oxygen and nitrogen gases do not react together. In the presence of very high temperatures nitrogen and oxygen do react together to form nitric oxide. These conditions are found in the combustion of coal and oil at electric power plants, and also during the **combustion of gasoline in automobiles.** Both of these sources contribute about equally to the formation of nitrogen oxides.

In areas of high automobile traffic, such as in large cities, the amount of nitrogen oxides emitted into the atmosphere can be quite significant. In the Los Angeles area, the main source of acid rain is from automobiles. In certain national parks such as Yosemite and Sequoia, automobile traffic is banned to limit the amount of air pollution damage to the trees and plants. This also has the effect of reducing the visual smog in the air.



Outside Links

- Clyde, Dale D. " Dynamite Demo? ." *J. Chem. Educ.* **1995** *72* 1130.
- Driscoll, Jerry A. *"* Acid Rain Demonstration: The Formation of Nitrogen Oxides as a By-Product of High-Temperature Flames in Connection with Internal Combustion Engines *."* *J. Chem. Educ.* **1997** *74* 1424.
- Foster, Natalie I.; Heindel, Ned D. "The discovery of nitroglycerine: Its preparation and therapeutic utility (SBS)." *J. Chem. Educ.* **1981**, *58*, 364.

- Leenson, I. A. " Approaching Equilibrium in the N_2O_4-NO_2 System: A Common Mistake in Textbooks." *J. Chem. Educ.* **2000** *77* 1652.

Contributors

- Charles Ophardt, Professor Emeritus, Elmhurst College, Virtual Chembook

2.2.1.6: Sources of Nitrogen Oxides is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2.1.7: Sources of Sulfur Oxides

It has been estimated that on a global basis, natural sources, such as volcanoes, contribute about as the same amount of sulfur oxides to the atmosphere as human industrial activities. This amounts to 75-100 million tons from each source per year. However, in industrial countries such as in Europe and North America, human activities contribute 95 % of the sulfur oxides and natural sources only 5 %. In the Western States, natural sources of sulfur oxides may be more important.

Human Sources of Sulfur Oxides

In 1980, emissions of sulfur dioxide totaled 24.1 million tons in the United States. Of this total 66 % came from electric power companies. Electric power companies that burn coal are a major source of sulfur oxides. Other industrial plants contributed about 22 %. Smelting of metals such as copper, zinc, lead, and nickel can produce large amounts of sulfur dioxide. In Canada, 45% of the emissions are from smelting operations, compared to only 6 % in the United States.

Coal contains mainly carbon with some hydrogen. When coal is burned it reacts with oxygen in the air to produce carbon dioxide and water and large amounts of heat.

$$C + O_2 \rightarrow CO_2 \tag{2.2.1.7.1}$$

In addition, coal may contain from 1-4 % of the element, sulfur. When the coal is burned with oxygen in the air, the sulfur is reacted to form **sulfur dioxide**.

$$S + O_2 \rightarrow SO_2 \tag{2.2.1.7.2}$$

Wood Smoke

In certain resort towns, a significant source of visible smog conditions results from the burning of large quantities of wood in fireplaces and stoves. The smoke contains solid particles which may provide the initial bit of solid or catalyst that initiates the reactions to produce sulfuric acid or nitric acid in the water droplets. This is a well recognized problem in Aspen and Vail Colorado. Steps are being taken to reduce the burning of wood.

Contributors

- Charles Ophardt, Professor Emeritus, Elmhurst College, Virtual Chembook

2.2.1.7: Sources of Sulfur Oxides is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

SECTION OVERVIEW

2.2.2: Aquatic Chemistry

Water, a natural occurring and abundant substance that exists in solid, liquid, and gas forms on the planet Earth, has attracted the attention of artists, engineers, poets, writers, philosophers, environmentalists, scientists, and politicians. Every aspect of life involves water as food, as a medium in which to live, or as the essential ingredient of life. The food-science aspects of water range from agriculture, aquaculture, biology, biochemistry, cookery, microbiology, nutrition, photosynthesis, power generation, to zoology. Even in the narrow sense of food technology, water is intimately involved in the production, washing, preparation, manufacture, cooling, drying, and hydration of food. Water is eaten, absorbed, transported, and utilized by cells. Facts and data about water are abundant and diverse. This article can only selectively present some fundamental characteristics of water molecules and their collective properties for readers when they ponder food science at the molecular level.

2.2.2.1: 1. Acid-Base Chemistry of Natural Aquatic Systems

2.2.2.2: 2. Carbonate Equilibria in Natural Waters

2.2.2.3: 3. Redox Equilibria in Natural Waters

2.2.2.4: 4. Solids in Contact With Natural Waters

2.2.2.5: Fundamental Characteristics of Water

2.2.2.6: Natural Water

2.2.2.7: Water Biology

2.2.2.8: Water Chemistry

2.2.2.9: Water Physics

2.2.2.10: Water Treatment

This page titled 2.2.2: Aquatic Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.2.2.1: 1. Acid-Base Chemistry of Natural Aquatic Systems

Natural waters contain a wide variety of solutes that act together to determine the pH, which typically ranges from 6 to 9. Some of the major processes that affect the acid-base balance of natural systems are:

- Contact with atmospheric carbon dioxide
- Input of acidic gases from volcanic and industrial emissions
- Contact with minerals, rocks, and clays
- Presence of buffer systems such as carbonate, phosphate, silicate, and borate
- Presence of acidic anions, such as $Fe(H_2O)_6^{3+}$
- Input and removal of CO_2 through respiration and photosynthesis
- Other biological processes, such as oxidation ($O_2 + H^+ + e^- \rightarrow H_2O$), nitrification, denitrification, and sulfate reduction.

In this chapter and also in the next one which deals specifically with the carbonate system, we will consider acid-base equilibria as they apply to natural waters. We will assume that you are already familiar with such fundamentals as the Arrhenius and Brfinsted concepts of acids and bases and the pH scale. You should also have some familiarity with the concepts of free energy and activity. The treatment of equilibrium calculations will likely extend somewhat beyond what you encountered in your General Chemistry course, and considerable emphasis will be placed on graphical methods of estimating equilibrium concentrations of various species.

Contributors and Attributions

Stephen Lower, Professor Emeritus (Simon Fraser U.) Chem1 Virtual Textbook

This page titled 2.2.2.1: 1. Acid-Base Chemistry of Natural Aquatic Systems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.2.2.2: 2. Carbonate Equilibria in Natural Waters

http://www.chem1.com/acad/webtext/pdf/c3carb.pdf

This page titled 2.2.2.2: 2. Carbonate Equilibria in Natural Waters is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.2.2.3: 3. Redox Equilibria in Natural Waters

http://www.chem1.com/acad/webtext/pdf/c3redox.pdf

Back to Advanced Aquatic Chemistry

Back to chem1 (Stephen Lower/Simon Frasier University)

This page titled 2.2.2.3: 3. Redox Equilibria in Natural Waters is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.2.2.4: 4. Solids in Contact With Natural Waters

http://www.chem1.com/acad/webtext/pdf/c3solids.pdf

This page titled 2.2.2.4: 4. Solids in Contact With Natural Waters is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stephen Lower via source content that was edited to the style and standards of the LibreTexts platform.

2.2.2.5: Fundamental Characteristics of Water

I. Introduction

Water, a natural occurring and abundant substance that exists in solid, liquid, and gas forms on the planet Earth, has attracted the attention of artists, engineers, poets, writers, philosophers, environmentalists, scientists, and politicians. Every aspect of life involves water as food, as a medium in which to live, or as the essential ingredient of life. The food-science aspects of water range from agriculture, aquaculture, biology, biochemistry, cookery, microbiology, nutrition, photosynthesis, power generation, to zoology. Even in the narrow sense of food technology, water is intimately involved in the production, washing, preparation, manufacture, cooling, drying, and hydration of food. Water is eaten, absorbed, transported, and utilized by cells. Facts and data about water are abundant and diverse. This article can only selectively present some fundamental characteristics of water molecules and their collective properties for readers when they ponder food science at the molecular level.

The physics and chemistry of water is the backbone of engineering and sciences. The basic data for the properties of pure water, which are found in the CRC Handbook of Chemistry and Physics (1), are useful for food scientists. However, water is a universal solvent, and natural waters contain dissolved substances present in the environment. All solutes in the dilute solutions modify the water properties. Lang's Handbook of Chemistry (2) gives solubilities of various gases and salts in water. Water usage in the food processing industry is briefly described in the Nalco Water Handbook (3). For water supplies and treatments, the Civil Engineering Handbook (4) provides practical guides. The Handbook of Drinking Water Quality (5) sets guidelines for waters used in food services and technologies. Wastewater from the food industry needs treatment, and the technology is usually dealt with in industrial chemistry (6). Most fresh food contains large amounts of water. Modifying the water content of foodstuffs to extend storage life and enhance quality is an important and widely used process (7).

A very broad view and deep insight on water can be found in "Water – A Matrix of Life" (8). Research leading to our present-day understanding of water has been reviewed in the series "Water – A Comprehensive Treatise" (9). The interaction of water with proteins (10, 11) is a topic in life science and food science. Water is the elixir of life and H_2O is a biomolecule.

II. Water and Food Technology

Water is an essential component of food (12). Philosophical conjectures abound as to how Earth evolved to provide the mantle, crust, atmosphere, hydrosphere, and life. Debates continue, but some scientists believe that primitive forms of life began to form in water (13). Complicated life forms developed, and their numbers grew. Evolution produced anaerobic, aerobic and photosynthetic organisms. The existence of abundant life forms enabled parasites to appear and utilize plants and other organisms. From water all life began (14). Homo sapiens are integral parts of the environment, and constant exchange of water unites our internal space with the external.

The proper amount of water is also the key for sustaining and maintaining a healthy life. Water transports nutrients and metabolic products throughout the body to balance cell contents and requirements. Water maintains biological activities of proteins, nucleotides, and carbohydrates, and participates in hydrolyses, condensations, and chemical reactions that are vital for life (15). On average, an adult consumes 2 to 3 L of water: 1-2 L as fluid, 1 L ingested with food, and 0.3 L from metabolism. Water is excreted via the kidney, skin, lung, and anus (16). The amount of water passing through us in our lifetimes is staggering.

Aside from minute amounts of minerals, food consists of plant and animal parts. Water is required for cultivating, processing, manufacturing, washing, cooking and digesting food. During or after eating, a drink, which consists of mostly water, is a must to hydrate or digest the food. Furthermore, water is required in the metabolic process.

Cells and living organisms require, contain and maintain a balance of water. An imbalance of water due to freezing, dehydration, exercise, overheating, etc. leads to the death of cells and eventually the whole body. Dehydration kills far more quickly than starvation. In the human body, water provides a medium for the transportation, digestion and metabolism of food in addition to many other physiological functions such as body temperature regulation (17).

Two-thirds of the body mass is water, and in most soft tissues, the contents can be as high as 99% (16). Water molecules interact with biomolecules intimately (9); it is part of us. Functions of water and biomolecules collectively manifest life. Water is also required for running households, making industrial goods, and generating electric power.

Water has shaped the landscape of Earth for trillions of years, and it covers 70% of the Earth's surface. Yet, for food production and technology it is a precious commodity. Problems with water supply can lead to disaster (5). Few brave souls accept the challenge to stay in areas with little rainfall. Yet, rainfall can be a blessing or a curse depending on the timing and amount. Praying for timely and bounty rainfall used to be performed by emperors and politicians, but water for food challenges scientists and engineers today.

III. Water Molecules and Their Microscopic Properties

Plato hypothesized four *primal substances*: water, fire, earth, and air. His doctrine suggested that a combination and permutation of various amounts of these four *primal substances* produced all the materials of the world. Scholars followed this doctrine for 2000 years, until it could not explain experimental results. The search of fundamental substances led to the discovery of hydrogen, oxygen, nitrogen, etc., as chemical elements. Water is made up of hydrogen (H) and oxygen (O). Chemists use H_2O as the universal symbol for water. The molecular formula, H_2O, implies that a water molecule consists of two H atoms and one O atom. However, many people are confused with its other chemical names such as hydrogen oxide, dihydrogen oxide, dihydrogen monoxide, etc.

A. Isotopic Composition of Water

The discoveries of electrons, radioactivity, protons, and neutrons implied the existence of isotopes. Natural isotopes for all elements have been identified. Three isotopes of hydrogen are protium (^{1}H), deuterium (D, ^{2}D or ^{2}H), and radioactive tritium (T, ^{3}T or ^{3}H), and the three stable oxygen isotopes are ^{16}O, ^{17}O, and ^{18}O. The masses and abundances of these isotopes are given in Table 1. For radioactive isotopes, the half-lives are given.

Table 1. Isotopes of hydrogen and oxygen; and isotopic water molecules molar mass (amu), relative abundance (%) or half life

Isotopes of hydrogen			**Stable isotopes of oxygen**		
^{1}H	**^{2}D**	**^{3}T**	**^{16}O**	**^{17}O**	**^{18}O**
1.007825	2.00141018	3.0160493	15.9949146	16.9991315	17.9991604
99.985%	0.015%	12.33 years	99.762%	0.038%	0.200%
Isotopic water molecules molar mass (amu) and relative abundance (%, ppm or trace)					
$H_2{}^{16}O$	**$H_2{}^{18}O$**	**$H_2{}^{17}O$**	**$HD^{16}O$**	**$D_2{}^{16}O$**	**$HT^{16}O$**
18.010564 99.78%	20.014810 0.20%	19.014781 0.03%	19.00415 0.0149%	19.997737 0.022 ppm	20.018789 trace

Random combination of these isotopes gives rise to the various isotopic water molecules, the most abundant one being $^1H_2{}^{16}O$ (99.78%, its mass is 18.010564 atomic mass units (amu)). Water molecules with molecular masses about 19 and 20 are present at some fractions of a percent. Although $HD^{16}O$ (0.0149%) is much more abundant than $D_2{}^{16}O$ (heavy water, 0.022 part per million), D_2O can be concentrated and extracted from water. In the extraction process, HDO molecules are converted to D_2O due to isotopic exchange. Rather pure heavy water (D_2O) is produced on an industrial scale especially for its application in nuclear technology, which provides energy for the food industry.

A typical mass spectrum for water shows only mass-over-charge ratio of 18 and 17 respectively for H_2O^+ and OH^+ ions in the gas phase. Other species are too weak for detection, partly due to condensation of water in mass spectrometers.

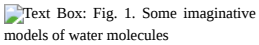
The isotopic composition of water depends on its source and age. Its study is linked to other sciences (18). For the isotopic analysis of hydrogen in water, the hydrogen is reduced to a hydrogen gas and then the mass spectrum of the gas is analyzed. For isotopes of oxygen, usually the oxygen in H_2O is allowed to exchange with CO_2, and then the isotopes of the CO_2 are analyzed. These analyses are performed on archeological food remains and unusual food samples in order to learn their origin, age, and history.

B. Structure and Bonding of Water Molecules

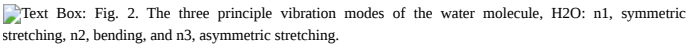
Chemical bonding is a force that binds atoms into a molecule. Thus, chemists use H-O-H or HOH to represent the bonding in water. Furthermore, spectroscopic studies revealed the H–O–H bond angle to be 104.5º and the H–O bond length to be 96

picometers (pm) for gas H_2O molecules (19). For solid and liquid, the values depend on the temperature and states of water. The bond length and bond angles are fundamental properties of a molecule. However, due to the vibration and rotational motions of the molecule, the measured values are average or equilibrium bond lengths and angles.

The mean van der Waals diameter of water has been reported as nearly identical with that of isoelectronic neon (282 pm). Some imaginary models of the water molecule are shown in Fig. 1


Text Box: Fig. 1. Some imaginative models of water molecules

An isolated water molecule is hardly static. It constantly undergoes a vibration motion that can be a combination of any or all of the three principle modes: symmetric stretching, asymmetric stretching, and bending (or deformation). These vibration modes are indicated in Fig. 2.


Text Box: Fig. 2. The three principle vibration modes of the water molecule, H2O: n1, symmetric stretching, n2, bending, and n3, asymmetric stretching.

Absorption of light (photons) excites water molecules to higher energy levels. Absorption of photons in the infrared (IR) region excites the vibration motion. Photons exciting the symmetric stretching, bending, and asymmetric stretching to the next higher energy levels have wave numbers 3656, 1594, and 3756 cm^{-1} respectively, for H_2O (20). These values and those for other water molecules involving only ^{16}O are given in Table 2

Table 2. Absorption frequencies of D2O, H2O, and HDO molecules for the excitation of fundamental modes to a higher energy level. Vibration mode

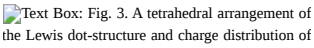
Symmetric stretching	3656	2726	2671
Bending	1594	1420	1178
Asymmetric stretching	3756	3703	2788

The spectrum of water depends on temperature and density of the gaseous H_2O. A typical IR spectrum for the excitation of only the fundamental vibration modes consists of three peaks around 1594, 3656 and 3756 cm^{-1}. Additional peaks due to excitation to mixed modes appear at higher wave numbers.

Rotating the H_2O molecule around the line bisecting the HOH angle by 180° (360°/2) results in the same figure. Thus, the molecules have a 2-fold rotation axis. There are two mirror planes of symmetry as well. The 2-fold rotation and mirror planes give the water molecules the symmetry point group C_{2v}.

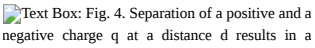
Rutherford alpha scattering experiment in 1909 showed that almost all atomic mass is in a very small atomic nucleus. In a neutral atom the number of protons in the nucleus is the same as the number of electrons around the nucleus. A proton and an electron have the same amount, but different kind of charge. Electrons occupy nearly all of the atomic volume, because the radius of an atom is 100,000 times that of the nucleus.

Electrons, in quantum mechanical view, are waves confined in atoms, and they exist in several energy states called orbitals. Electrons in atoms and molecules do not have fixed locations or orbits. Electron states in an element are called electronic configurations, and their designation for H and O are $1s^1$, and $1s^2 2s^2 2p^4$, respectively. The superscripts indicate the number of electrons in the orbitals 1*s*, 2*s* or 2*p*. The electronic configuration for the inert helium (He) is $1s^2$, and $1s^2$ is a stable core of electrons. Bonding or valence electrons are $1s^1$ and $2s^2 2p^4$ for H and O respectively.


Text Box: Fig. 3. A tetrahedral arrangement of the Lewis dot-structure and charge distribution of the H2O molecule.

The valence bond approach blends one 2*s* and three 2*p* orbitals into four bonding orbitals, two of which accommodate two electron pairs. The other two orbitals have only one electron each, and they accommodate electrons of the H atoms bonded to O, thus forming the two H–O bonds. An electron pair around each H atom and four electron pairs around the O atom contribute stable electronic configurations for H and O, respectively. The Lewis dot-structure, Fig. 3, represents this simple view. The two bonding and two lone pairs are asymmetrically distributed with major portions pointing to the vertices of a slightly distorted tetrahedron in 3-dimensional space. The two lone pairs mark slightly negative sites and the two H atoms are slightly positive. This charge distribution around a water molecule is very important in terms of its microscopic and macroscopic, chemical and physical properties described later. Of course, the study of water continues and so does the evolution of bonding theories. Moreover, the distribution of electrons in a single water molecule is different from those of dimers, clusters, and bulk water.

The asymmetric distribution of H atoms and electrons around the O atom results in positive and negative sites in the water molecule. Thus, water consists of polar molecules.

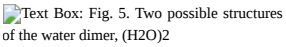

Text Box: Fig. 4. Separation of a positive and a negative charge q at a distance d results in a dipole, m = d q

The **dipole moment, *m*,** is a measure of polarity and a useful concept. A pair of opposite charge, *q*, separated by a distance, *d*, has a dipole moment of $m = d\ q$ with the direction pointing towards the positive charge as shown in Fig. 4.

The dipole moment of individual water molecules is $6.187\acute{}10^{-30}$ C m (or 1.855 D) (21). This quantity is the vector resultant of two dipole moments due to the O–H bonds. The bond angle H–O–H of water is 104.5°. Thus, the dipole moment of an O–H bond is $5.053\acute{}10^{-30}$ C m. The bond length between H and O is 0.10 nm, and the partial charge at the O and the H is therefore $q = 5.053\acute{}10^{-20}$ C, 32 % of the charge of an electron ($1.6022\acute{}10^{-19}$ C). Of course, the dipole moment may also be considered as separation of the electron and positive charge by a distance 0.031 nm.

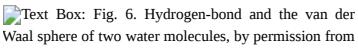
It should be pointed out that the dipole moments of liquid and solid water appear to be higher due to the influence of neighboring molecules. For the liquid and solid, macroscopic properties need be considered.

C. Hydrogen Bonds


Text Box: Fig. 5. Two possible structures of the water dimer, (H2O)2

Attraction among water molecules is more than polar-polar in nature. The O atoms are small and very electronegative. As a result, the positive H atoms (protons) are very attractive to the negative O atoms of neighboring molecules. This O- - -H – O strong attraction is called a **hydrogen bond**, a concept popularized by L. Pauling (22). Furthermore, hydrogen atoms bonded to atoms of N and F, neighboring elements of O in the periodic table, are positive, and they form hydrogen bonds with atoms of N, O, or F. The strength of hydrogen bonds depends on the X–H - - - Y (X or Y are N, O, or F atoms) distances and angles; the shorter the distances, the stronger are the hydrogen bonds.

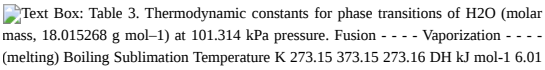
When two isolated water molecules approach each other, a dimer is formed due to hydrogen bonding. The dimer may have one or two hydrogen bonds. Dimers exist in gaseous and liquid water. When more water molecules are in close proximity, they form trimmers, tetramers and clusters. Hydrogen bonds are not static, they exchange protons and partners constantly. Hydrogen bonding is a prominent feature in the structures of various solid phases of water usually called ice as we shall see later.


Text Box: Fig. 6. Hydrogen-bond and the van der Waal sphere of two water molecules, by permission from M.F. Martin (23), sbu.ac.uk/water/molecule.html

Water molecules not only form hydrogen bonds among themselves, they form hydrogen bonds with any molecule that contains N–H, O–H and F–H bonds. Foodstuffs such as starch, cellulose, sugars, proteins, DNA, and alkaloids contain N–H and O–H groups, and these are both H-donors and H-acceptors of hydrogen bonds of the type N- - H–O, O- -H–N, N- -H–N, etc. A dimer depicting the hydrogen bond and the van der Waal sphere of two molecules is shown in Fig. 6 (23).

Carbohydrates (starch, cellulose and sugars) contain H–C–O–H groups. The O–H groups are similar to those of water molecules, and they are H-acceptors and H-donors for hydrogen bonds. Proteins contain O–H, R–NH_2 or R_2>NH groups, and the O–H and N–H groups are both H-donors and H-acceptors for the formation of hydrogen bonds. Thus, water molecules have intimate interactions with carbohydrates and proteins.

IV. Macroscopic Properties of Water


Text Box: Table 3. Thermodynamic constants for phase transitions of H2O (molar mass, 18.015268 g mol–1) at 101.314 kPa pressure. Fusion - - - - Vaporization - - - - (melting) Boiling Sublimation Temperature K 273.15 373.15 273.16 DH kJ mol-1 6.01 40.66 51.06 DS J mol-1 K-1 22.00 108.95 186.92 DE kJ mol-1 6.01 37.61 48.97 DV L mol-1 -1.621 30.10 - Energy change of phase transition: DH, enthalpy; DS, entropy; DE, internal energy. DV: volume change of phase transition.

Collectively, water molecules exist as gas, liquid, or solid depending on the temperature and pressure. These phases of water exhibit collective or macroscopic properties such as phase transitions, crystal structures, liquid structures, vapor pressures, and volume-pressure relationships of vapor. In addition, energies or enthalpies for melting, vaporization and heating are also important for applications in food technology.

Thermodynamic constants for phase transitions given in Table 3 are those of pure water. Natural waters, of course, contain dissolved air, carbon dioxide, organic substances, microorganisms, and minerals. Water in food or used during food processing usually contains various organic

and inorganic substances. These solutes modify the properties of water and caution should be taken to ensure proper values are applied in food technology.

The triple point and boiling points of water are defined as 273.16 and 373.16 K (kelvin) in the SI unit of temperature, respectively. Thus, the temperature differences can be in units of K or oC.

Water has many unusual properties due to its ability to form hydrogen bonds and its large dipole moment. As a result, the melting, boiling and critical points for water are very high compared to substances of similar molar masses. In general, the higher the molar masses, the higher are the melting and boiling points of the material. Associated with these properties are its very large heat of melting, heat capacity, heat of vaporization and heat of sublimation. Moreover, its surface tension and viscosity are also very large. Thermodynamic energies, and volume changes for phase transitions of H_2O are summarized in Table 3. These data are mostly taken from the Encyclopedia of Chemical Technology, Vol. 25 (1991) (24).

A. Crystal Structures and Properties of Ice

Text Box: Table 4. Properties of Ice at 273.15 K Heat of formation DH f 292.72 kJ mol-1 Density 0.9168 g cm-3 Heat capacity 2.06 J g-1 K-1 DHfusion 6.01 kJ mol-1 Dielectric constant at 3 kHz 79 Thermal expansion coefficient Volumetric 120´10-6 cm3 g-1 K-1 Linear 52.7´10-6 cm g-1 K-1

Hydrogen bonding is prominent in the crystal structures of various solid phases of H_2O. The triple point of water is at 273.16 K and 4.58 torr (611 Pa). The melting point at 1.00 atm (760 torr or 101.325 kPa) is used to define the Kelvin scale as 273.15 K. When water freezes at these temperatures and atmospheric pressure or lower, the solids are hexagonal ice crystals usually designated as I_h. Properties of I_h are given in Table 4. Snowflakes have many shapes because their growth habit depends on temperature and vapor pressure, but they all exhibit hexagonal symmetry, due to the hexagonal structure of ice (25).

However, from a geometric point of view, the same bonding may also be arranged to have cubic symmetry. The existence of cubic ice has been confirmed. When water vapor deposits onto a very cold, 130 – 150 K, surface or when small droplets are cold under low pressure at high altitude, the ice has a cubic symmetry usually designated as I_c. At still higher pressures, different crystal forms designated as ice II, III, IV, … etc., up to 13 phases of cubic, hexagonal, tetragonal, monoclinic, and orthorhombic symmetries have been identified (26). The polymorphism of solid water is very complicated. Some of these ice forms are made under very high pressures, and water crystallizes into solid at temperatures above the normal melting or even boiling temperatures. Ice VII is formed above 10 G Pa (gigapascal) at 700 K (26).

When liquid water is frozen rapidly, the molecules have little chance of arranging into crystalline ice. The frozen liquid is called **amorphous ice** or **glassy ice**.

Text Box: Fig. 7. Ordered crystal structure of ice, bond distance: O–H, 100 pm; H- - -O, 175 pm; and O–H- - -O, 275 pm.

The basic relationships between nearest neighboring water molecules are the same in both I_h and I_c. All O atoms are bonded to four other O atoms by hydrogen bonds, which extend from an oxygen atom towards the vertices of a tetrahedron. A sketch of the crystal structure of hexagonal I_h is shown in Fig. 7 (27). In I_h, hydrogen positions are somewhat random due to thermal motion, disorder and exchanges. For example, the hydrogen may shift between locations to form H_3O^+ and OH^- ions dynamically throughout the structure. In this structure, bond angles or hydrogen-bond angles around oxygen atoms are those of the idealized tetrahedral arrangement of 109.5^o rather than 104.5^o observed for isolated molecules. Formation of the hydrogen bond in ice lengthens the O–H bond distance, 100 pm compared to 96 pm in a single water molecule. The diagram illustrates a crystal structure that is completely hydrogen bonded, except for the molecules at the surface.

Each O atom of hexagonal ice I_h is surrounded by four almost linear O–H - - O hydrogen bonds of length 275 pm, in a tetrahedral fashion. Each C atom of cubic diamond is also surrounded by four C–C covalent bonds of length 154 pm. Thus, the tetrahedral coordination can either be cubic or hexagonal, from a geometrical viewpoint. Indeed, the uncommon cubic ice and hexagonal diamond have been observed, giving a close relationship in terms of spatial arrangement of atoms between ice and diamond (26). Strong hydrogen bonds make ice hard, but brittle. The structure is related to its physical properties, which vary with temperature.

The pressure of H_2O vapor in equilibrium with ice is called the **vapor pressure of ice,** which decreases as the temperature decreases. At the triple point or 0° C, the pressure is 611.15 Pa. When ice is slightly overheated to 0.01° C, the pressure increases to 611.657 Pa. However, at this temperature, the vapor pressure of liquid water is lower. The vapor pressures of ice between 0° C and – 40° C are listed in Table 5 at 1°C interval. Various models can be used to estimate the vapor pressure at other temperatures. One

Text Box: Table 5. Aqueous vapor pressure (Pa) in equilibrium with ice between 0 and –40°C (at 0.01°C, p is 611.657 Pa) t (° C) – 0< – 1< –2< –3< – 4< – <0 611.15 259.90 103.26 38.01 12.84 – <1 562.67 237.74 93.77 34.24 – <2 517.72 217.32 85.10 30.82 – <3 476.06 198.52 77.16 27.71 – <4 437.47 181.22 69.91 24.90 – <5 401.76 165.30 63.29 22.35 7.20 – <6 368.73 150.68 57.25 20.04 – <7 338.19 137.25 51.74 17.96 – <8 309.98 124.92 46.73 16.07 – <9 283.94 113.62 42.16 14.37 Missing digits of t, <, in the row are found in the column, and vice versa. These values are taken from the CRC Handbook of Chemistry and Physics (1).

method uses the Clausius-Clapeyron differential equation

$$\frac{dp}{dT} = \frac{H}{T\,DV}$$

where p is the pressure, T is the temperature (K), H is the latent heat or **enthalpy** of phase transition, and DV is the difference in volume of the phases. The enthalpy of sublimation for ice depends on the temperature. At the freezing point, the enthalpy of sublimation for ice is 51 (51.06 in Table 3) kJ mol^{-1}, estimated from the vapor pressure at 0 and –1° C. The enthalpy of sublimation is required to overcome hydrogen bonding, dipole, and intermolecular attractions. The energy required in freeze-drying processes varies, depending on temperature and other conditions. Water in solutions and in food freezes below 0°C.

The number of hydrogen bonds is twice the number of water molecules, when surface water molecules are ignored. The energy required to separate water molecules from the solid is the enthalpy of sublimation (55.71 J mol^{-1}). Half of this value, 26 kJ mol^{-1}, is the energy to separate the H- -O linkages, and it translates into 0.26 eV, per H- -O bond. These values are close to those obtained by other means (25, 26, 28, 29, 30). Several factors contribute to this linkage, and the hydrogen-bond energy is less than 0.26 eV.

Text Box: Table 6 Properties of liquid water at 298 K Heat of formation DHf 285.89 kJ mol-1 Density at 3.98o C 1.000 g cm-3 Density at 25o C 0.9970480 g cm-3 Heat capacity 4.17856 J g-1 K-1 DHvaporization 55.71 kJ mol-1 Dielectric constant 80. Dipole moment 6.24´10-30 C m Viscosity 0.8949 mP? s–1 Velocity of sound 1496.3 m s–1 Volumetric thermal expansion coefficient 0.0035 cm3 g-1 K-1

B. Properties of Liquid Water

The macroscopic physical properties of this common but eccentric fluid at 298 K (25° C) are given in Table 6. Water has an unusually high melting and boiling points for a substance of molar mass of only 18 daltons. Strong hydrogen bonds and high polarity accounts for this.

The heat of formation is the energy released when a mole of hydrogen and half a mole of oxygen at 298 K and 1.00 atm react to give one mole of water at 298 K. This value differs from that for ice in Table 4 due to both temperature and phase differences. As temperature increases, the average kinetic energy of molecules increases, and this affects water's physical properties. For example, surface tension of water decreases, whereas the thermal conductance increases as the temperature increases. Heat capacity at constant pressure (Cp), vapor pressure, viscosity, thermal conductance, dielectric constant, and surface tension in the temperature range 273–373 K (0–100°C) are given in Table 7.

Table 7. Properties of liquid water in the range 273 – 373 K (0 – 100° C)

Temp. Heat Viscosity Thermal Dielectric Surface

t capacity *C*p conductance constant tension

(°C) (J g^{-1} K^{-1}) (mPa s) (W K^{-1} m^{-1}) (mN m^{-1})

0 4.2176 1.793 561.0 87.90 75.64

10 4.1921 1.307 580.0 83.96 74.23

20 4.1818 1.002 598.4 80.20 72.75

30 4.1784 0.797 615.4 76.60 71.20

40 4.1785 0.653 630.5 73.17 69.60

50 4.1806 0.547 643.5 69.88 67.94

60 4.1843 0.466 654.3 66.73 66.24

70 4.1895 0.404 663.1 63.73 64.47

80 4.1963 0.354 670.0 60.86 62.67

90 4.2050 0.315 675.3 58.12 60.82

100 4.2159 0.282 679.1 55.51 58.91

More detailed data can be found in the CRC Handbook of Chemistry and Physics (1)

Liquid water has the largest heat capacity per unit mass of all substances. Large quantities of energy are absorbed or release when its temperatures changes. The large heat capacity makes water an excellent reservoir and transporter of energy. A large body of water moderates climate. The heat capacity C_p of water varies between 4.1 to 4.2 J g^{-1} K^{-1} (74 to 76 J mol^{-1} K^{-1}) even at temperature above 100° C and high pressure. The enthalpy of vaporization for water is also very large (55.71 kJ mol^{-1} at 298 K). Thus, energy consumption is high for food processing when water is involved.

Water and aqueous solutions containing only low molar-mass solutes are typical Newtonian fluids for which the **shear stress** is proportional to **shear strain rate**. Viscosity is the ratio of shear stress to shear strain rate. On the other hand, viscosity of solutions containing high molar-mass substances depends on shear strain rate. For pure water, the viscosity decreases from 1.793 to 0.282 mPa s (millipascal seconds; identical to centipoise (cp)) as temperature increases from 0 to 100° C. Thus, the flow rate through pipes increases as water or solution temperature increases.

The dielectric constant of water is very large, and this enables water to separate ions of electrolytes, because it reduces the electrostatic attraction between positive and negative ions. Many salts dissolve in water. When an electric field is applied to water, its dipole molecules orient themselves to decrease the field strength. Thus, its dielectric constant is very large. The dielectric constant decreases as temperature increases, because the percentage of molecules involved in hydrogen bonding and the degree of order decrease (28, 29). The measured dielectric constant also depends on the frequency of the applied electric field used in the measurement, but the variation is small when the frequency of the electric field is less than 100 MHz. The dielectric behavior of water allows water vapor pressure to be sensed by capacitance changes when moisture is absorbed by a substance that lies between the plates of a capacitor. These sensors have been developed for water activity measurement (31).

The light absorption coefficients are high in the infrared and ultraviolet regions, but very low in the visible region. Thus, water is transparent to human vision.

Text Box: Table 8. Vapor pressure (kPa) of liquid H2O between triple and critical points at every 10° C t ° C 0<< 1<< 2<< 3<< <00 0.6113 101.32 1553.6 8583.8 <10 1.2281 143.24 1906.2 9860.5 <20 2.3388 198.48 2317.8 11279 <30 4.2455 270.02 2795.1 12852 <40 7.3814 361.19 3344.7 14594 <50 12.344 475.72 3973.6 16521 <60 19.932 617.66 4689.4 18665 <70 31.176 791.47 5499.9 21030 <80 47.375 1001.9 6413.2 22055* <90 70.117 1254.2 7438.0 * Critical pressure at 373.98° C Missing digits of t, <, in the row are found in the column, and vice versa. Values from the CRC Handbook of Physics and Chemistry (2003) (1), which lists vapor pressure at 1o interval.

The variation of vapor pressure as a function of temperature is the bases for defining **water activities** of food. Liquid water exists between the triple-point and the critical-point temperatures (0 – 373.98°C) at pressures above the vapor pressures in this range.

As with ice, the vapor pressure of liquid water increases as the temperature increases. Vapor pressures of water (in kPa instead of Pa for ice in Table 5) between the triple and critical points, at 10°C interval, are given in Table 8. When the vapor pressure is 1.00 atm (101.32 kPa) the temperature is the boiling point (100° C). At slightly below 221°C, the vapor pressure is 2.00 atm. The critical pressure at the critical temperature, 373.98° C, is 217.67 atm (22,055 kPa). Above this temperature, water cannot be liquefied, and the phase is called **supercritical water**.

The partial pressure of H_2O in the air at any temperature is the **absolute humidity**. When the air is saturated with water vapor, the **relative humidity** is 100%. The unsaturated vapor pressure divided by the vapor pressure of water as given in Table 8 at the temperature of the air is the **relative humidity**. The temperature at which the vapor pressure in the air becomes saturated is the **dew point**, at which dew begins to form. However when the dew point is below 273 K or 0ºC, ice crystals (frost) begin to form. Thus, the relative humidity can be measured by finding the dew point. Dividing the vapor pressure at the dew point by the vapor pressure of water at the temperature of the air gives the relative humidity. The transformations between solid, liquid, and gaseous water play important roles in hydrology and in the transformation of the environment on Earth. Phase transitions of water combined with the energy from the sun make the weather.

Density is a collective property, and it varies with temperature, isotopic composition, purity, etc. The International Union of Pure and Applied Chemistry (IUPAC) has adopted the density of pure water from the ocean as the density standard. The isotopic

Text Box: Table 9. The density of water (g / mL) as a function of temperature between 0 and 39o C (31). t ° C 0< 1< 2< 3< <0 0.9998426 0.9997021 0.9982063 0.9956511 <1 0.9999015 0.9996074 0.9979948 0.9953450 <2 0.9999429 0.9994996 0.9977730 0.9950302 <3 0.9999627 0.9993792 0.9975412 0.9947971 <4 0.9999750* 0.9992464 0.9972994 0.9943756 <5 0.9999668 0.9991016 0.9970480 0.9940359 <6 0.9999430 0.9989450 0.9967870 0.9936883 <7 0.9999043 0.9987769 0.9975166 0.9933328 <8 0.9998509 0.9985976 0.9962370 0.9929695 <9 0.9997834 0.9984073 0.9956511 0.9925987 Missing digits of t, <, in the row are found in the column, and vice versa.

composition of ordinary water is constant, and the density of pure water between 0 and 39^o C extracted from (32) is given in Table 9.

The density of cage-like ice I_h, due to 100% of its molecules involved in hydrogen-bonded is only 9% lower than that of water. This indicates that water has a high percentage of molecules involved in the transient and dynamic hydrogen bonding. The percentage of hydrogen-bonded water molecules in water decreases as temperature increases, causing water density to increase. As temperature increases the thermal expansion causes its density to decreases. The two effects cause water density to increase from 0 to 3.98^{o}C, reaching its maximum of 1.0000 g mL^{-1} and then decrease as temperature increases.

Incidentally at 8^o C, the density of water is about the same as that at 0^o C. At 25^o C, the density decreases 0.3 % with respect to its maximum density, whereas at 100^o C, it decreases by 4 %. Dense water sinks, and convection takes place when temperature fluctuates at the surface of lakes and ponds, bringing dissolved air and nutrients to various depths of waters for the organism living in them. On the other hand, the pattern of density dependence on temperature of water makes temperatures at the bottoms of lakes and oceans vary little if the water is undisturbed. When water freezes, ice begins to form at the surface, leaving the water at some depth undisturbed. Water at the bottom remains at 4^{o}C preserving various creatures living in water.

When hydrogen-bonded to tissues and cells or in food, water has a unique order and structure, and the vapor pressure and density differ from those of pure water. Yet, the collective behavior of water molecules sheds some light regarding their properties in food, cells, tissues, and solutions.

V. Chemical Properties of Water

Water is a chemical as is any substance, despite the confusion and distrust of the public regarding the term “chemical”. Thus, water has lots of interesting chemical properties. It interacts intimately with components of food particularly as a solvent, due to its dipole moment and its tendency to form hydrogen bond. These interactions affect the chemical properties of nutrients, including their tendency to undergo oxidation or reduction, to act as acids or bases, and to ionize.

A. Water as a Universal Solvent

Water is dubbed a universal solvent, because it dissolves many substances due to strong interactions between water molecules and those of other substances. Entropy is another driving force for a liquid to dissolve or mix with other substances. Mixing increases disorder or entropy.

(a) hydrophobic effect and hydrophilic effect

Because of its large dielectric constant, high dipole moment and ability to donate and accept protons for hydrogen bonding, water is an excellent solvent for polar substances and electrolytes, which consist of ions. Molecules strongly interact with or love water molecules are **hydrophilic**, due to hydrogen bonding, polar-ionic or polar-polar attractions. Nonpolar molecules that do not mix with water are **hydrophobic** or **lipophilic**, because they tend to dissolve in oil. Large molecules such as proteins and fatty acids that have hydrophilic and hydrophobic portions are **amphipathic** or **amphiphilic.** Water molecules strongly intermingle with hydrophilic portions by means of dipole-dipole interaction or hydrogen bonding.

The lack of strong interactions between water molecules and lipophilic molecules or the nonpolar portions of amphipathic molecules is called the **hydrophobic effect**, a term coined by Charles Tanford (33). Instead of a direct interaction with such solutes, water molecules tend to form hydrogen-bonded cages around small nonpolar molecules when the latter are dispersed into water. Hydrogen-bonded water molecules form cages, called **hydrates** or **clathrates**, For example, the clathrate of methane forms stable crystals at low temperatures (34).

Nonpolar chains in proteins prefer to stay together as they avoid contact with water molecules. Hydrophilic and hydrophobic effects play important roles for the stability and state of large molecules such as enzymes, proteins, and lipids. Hydrophobic portions of these molecules stay together forming pockets in globular proteins. Hydrophilic and hydrophobic effects cause nonpolar portions of phospholipids, proteins, and cholesterol to assemble into bilayers or biological membranes (34).

(b) Hydration of ions

Text Box: Fig. 8. The first hydration sphere of most cations M(H2O)6+, and anions X(H2O)6–1 . Small water molecules are below the plane containing the ions, and large water molecules are above the plane.

Due to its high dielectric constant, water reduces the attractions among positive and negative ions of electrolytes and dissolves them. The polar water molecules coordinate around ions forming hydrated ions such as $Na(H_2O)_6^+$, $Ca(H_2O)_8^{2+}$, $Al(H_2O)_6^{3+}$ etc. Six to eight water molecules form the first sphere of hydration around these ions. Fig. 8 is a sketch of the interactions of water molecules with ions. The water molecules point their negative ends of their dipoles towards positive ions, and their positive ends towards negative ions. Molecules in the hydration sphere constantly and dynamically exchange with those around them. The number of hydrated-water molecules and their lifetimes have been studied by various methods. These studies reveal that the hydration sphere is one-layer deep, and the life times of these hydrated-water molecules are in the order of picoseconds (10^{-12} s). The larger negative ions also interact with the polar water molecules, not as strong as those of cations. The presence of ions in the solution changes the ordering of molecules even if they are not in the first hydration sphere (9).

The hydration of ions releases energy, but breaking up ions from a solid requires energy. The amounts of energy depends on the substance, and for this reason, some are more soluble than others. Natural waters in the ocean, streams, rivers and lakes are in contact with minerals and salts. The concentrations of various ions depend on the solubility of salts (35) and the contact time.

Drinking water includes all waters used in growth, processing, and manufacturing of food. J. De Zuane divides ions in natural water into four types in The Handbook on Drinking Water Quality (5).

Type A includes arsenic, barium, cadmium, chromium, copper, fluoride, mercury, nitrate, nitrite, and selenium ions. They are highly toxic, yet abundant.

Type B includes aluminum, nickel, sodium, cyanide, silver, zinc, molybdenum, and sulfate ions. Their concentrations are also high, but they are not very toxic.

Type C consists of calcium, carbonate, chloride, iron, lithium, magnesium, manganese, oxygen, phosphate, potassium, silica, bromine, chlorine, and iodine and ozone. They are usually present at reasonable levels.

Type D ions are present usually at low levels: antimony, beryllium, cobalt, tin, thorium, vanadium and thallium.

Most metals are usually present in water as cations, with a few as anions. However, some chemical analyses may not distinguish their state in water. The most common anions are chloride, sulfate, carbonate, bicarbonate, phosphate, bromide, iodide, etc. Toxicity is a concern for ions in water, but some of these ions are essential for humans.

Pure water has a very low electric conductivity, but ions in solutions move in an electric field making electrolyte solutions highly conductive. The conductivity is related to total dissolved solids (TDS), salts of carbonate, bicarbonate, chloride, sulfate, and nitrate. Sodium, potassium, calcium and magnesium ions are often present in natural waters because their soluble salts are common minerals in the environment. The solubilities of clay (alumina), silicates, and most common minerals in the Earth crust, are low.

(c) Hard waters and their treatments

Waters containing plenty of dissolved CO_2 (H_2CO_3) are acidic and they dissolve $CaCO_3$ and $MgCO_3$. Waters with dissolved Ca^{2+}, Mg^{2+}, HCO_3^- and CO_3^{2-} are called **temporary hard waters** as the hardness can be removed by boiling, which reduces the solubility of CO_2. When CO_2 is driven off, the solution becomes less acidic due to the following equilibria (the double arrows, D, indicate reversible reactions):

H^+ (aq) + HCO_3^- (aq) D H_2CO_3 (aq) D H_2O + CO_2 (g)

HCO_3^- (aq) D H^+ (aq) + CO_3^{2-} (aq)

Reducing the acidity increases the concentration of CO_3^{2-} and solids $CaCO_3$ and $MgCO_3$ precipitate:

Ca^{2+} (aq) + CO_3^{2-} (aq) D $CaCO_3$ (s)

Mg^{2+} (aq) + CO_3^{2-} (aq) D $MgCO_3$ (s)

Water containing less than 50 mg L^{-1} of these substances is considered soft; 50-150 mg L^{-1} moderately hard; 150 - 300 mg L^{-1} hard; and more than 300 mg L^{-1} very hard.

For the **lime treatment**, we determine the amount of dissolved Ca^{2+} and Mg^{2+} first; and then add an equal number of moles of lime, $Ca(OH)_2$, to remove them by these reactions:

$$Mg^{2+} + Ca(OH)_2(s) \rightarrow Mg(OH)_2(s) + Ca^{2+}$$

$$Ca^{2+} + 2\ HCO_3^- + Ca(OH)_2(s) \rightarrow 2\ CaCO_3(s) + 2\ H_2O$$

Permanent hard waters contain sulfate (SO_4^{2-}) ions with Ca^{2+} and Mg^{2+}. Calcium ions, Ca^{2+}, of the sulfate solution can be removed by adding sodium carbonate:

$$Ca^{2+} + Na_2CO_3 \rightarrow CaCO_3\ (s) + 2\ Na^+$$

Hard waters cause scales or deposits to build up in boilers and pipes, and they are usually softened by ion exchange with resins or zeolites. In these processes, the calcium and magnesium ions are taken up by the zeolite or resin that releases sodium or hydrogen ions into the water. Reverse osmosis has also being used to soften hard water.

However, water softening replaces desirable calcium and other ions by sodium ions. Thus, soft waters are not suitable drinking waters. Incidentally, bakers use hard water because the calcium ions strengthen the gluten proteins in dough mixing. Some calcium salts are added to dough to enhance bread quality.

(d) Properties of aqueous solutions

Waters containing dissolved substances are aqueous solutions; their physical properties differ from those of pure water. For example, at the same temperature, the H_2O vapor pressures of solutions are lower than that of pure water, resulting in boiling point elevation (higher), freezing point depression (lower) and osmotic pressure.

Several ways can be used to express concentrations: part per million (ppm), percent, moles per liter, mole fraction, etc. The mole fraction of water is the fraction of water molecules among all molecules and ions in the system. The vapor pressure of an ideal solution, $P_{solution}$, is the vapor pressure of water (at a given temperature), P°_{water}, modified by the mole fraction X_{water}.

$P_{solution} = X_{water}\ P^{\circ}_{water.}$ ($X_{water} < 1$)
If the solute has a significant vapor pressure, P_{solute} is also modified by its mole fraction,

$P_{solute} = X_{solute}\ P^{\circ}_{solute.}$

For non-ideal solutions, in which water and solute strongly interact, the formulas require modifications. A practical method is to use an **effective mole fraction** X defined by:
$P_{solution} / P^{\circ}_{water.} = X$

In any case, the vapor pressures of solutions containing nonvolatile electrolytes are lower than those of pure water at their corresponding temperature.

Phase transitions take place when the vapor pressures of the two phases are the same. Because solutions' vapor pressures are lower, their melting points are lower but their boiling points are higher. The difference in temperature, DT, is proportional to the concentrations of all solutes, $m_{all\text{-}solute}$ (molality),

$\Delta T = K\ m_{all\text{-}solute}$
where K is either the molar boiling point elevation constant, K_b, or the molar freezing point depression constant K_f. For water, K_f = 1.86 K L kg^{-1}, and K_b = 0.52 K L kg^{-1}. Due to ionization of electrolytes, positive and negative ions should be treated as separate species and all species should be included in $m_{all\text{-}solute}$.

The tendency of water molecules from a dilute solution to diffuse into a more concentrated solution, through semipermeable membranes, has a measurable quantity called **osmotic pressure**, p, which is proportional to the concentration (mol per kg of water) of all dissolved species, $m_{all\text{-}solute}$ in mol kg^{-1} and temperature T in K,

$\Pi = -\ m_{all\text{-}solute}\ R\ T$
where R is the gas constant 8.3145 J mol^{-1} K^{-1}. Water molecules diffuse from pure water (Π = 0) into the solution, and the osmotic pressure is therefore given as a negative value here. Theoretically, a solution with $m_{all\text{-}solute}$ = 1.0 mol kg^{-1} of water, Π = – 2477 J kg or
– 2.477 kJ kg at 298 K. Note $m_{all\text{-}solute} = m\ i$; ($i$ being the number of ions produced by the solute) in the van't Hoff equation, which is often used in other literature.

Solutions having the identical osmotic pressure are isotonic. Applying more pressure to a solution to compensate for the osmotic pressure causes water molecules to diffuse through membranes, generating pure or fresh waters. This process is called **reverse**

osmosis, and it has been used to soften water or desalinate seawater, converting it to fresh water.

The lowering of vapor pressure and the osmotic pressure of solutions play important roles in hydration and dehydration of food and in living cells. Solutions containing proper concentrations of nutrients and electrolyte have been used to medically treat dehydrated patients. J.R. Cade and his coworkers applied these principles to formulate drinks for athletes; he and his coworkers were credited as the inventor of the sports drink Gaterade (36). The concept of a balanced solution for hydration became a great business decades after its invention.

B. Acidity and Alkalinity of Water

Acidity and alkalinity are also important characteristics of water due to its dynamic self-ionization equilibrium,
H_2O (l) D $H^+ + OH^-$, $K_w = [H^+][OH^-] = 1 \acute{} 10^{-14}$ at 298 K and 1 atm
where $[H^+]$ and $[OH^-]$ represent the molar concentrations of H^+ (or H_3O^+) and OH^- ions, respectively, and K_w is called the **ion product of water**, (see tables in ref. 1 and 36). Values of K_w under various conditions have been evaluated theoretically (37, 38). Solutions in which $[H^+] = [OH^-]$ are said to be neutral. At 298 K, for a neutral solution,
pH = – log [H+] = pOH = – log $[OH^-]$ = 7 (at 298 K)

The H^+ ions or protons dynamically exchange with protons in other water molecules. The self-ionization and equilibrium are present in all aqueous solutions, including acid and base solutions, as well as in pure water. Water is both an acid and a base.

Strong acids such as $HClO_4$, $HClO_3$, HCl, HNO_3, and H_2SO_4 ionize completely in their solutions to give H^+ (H_3O^+) ions and anions: ClO_4^-, ClO_3^-, Cl^-, NO_3^-, and HSO_4^-, respectively. Strong bases such as NaOH, KOH, and $Ca(OH)_2$ also ionize completely giving OH^- ions and Na^+, K^+, and Ca^{2+} ions respectively. In an acidic solution, $[H^+]$ is greater than $[OH^-]$. In a 1.0 mol L^{-1} HCl solution, $[H^+]$ = 1.0 mol L^{-1}, pH = 0.

Weak acids such as formic acid, HCOOH, acetic acid (CH_3COOH), ascorbic acid ($C_6H_8O_6$), oxalic acid ($H_2C_2O_4$), carbonic acid (H_2CO_3), benzoic acid (C_6H_5COOH), malic acid ($C_4H_6O_5$), lactic acid $H_3CCH(OH)COOH$, and phosphoric acid (H_3PO_4) also ionize in their aqueous solutions, but not completely. The ionization of acetic acid, is represented by the equilibrium,
CH_3COOH (aq) D H^+ (aq) + CH_3COO^- (aq), $K_a = 1.75 \times 10^{-5}$ at 298 K
where K_a is the acid dissociation constant.

The solubility of CO_2 in water increases with its (CO_2 partial) pressure, according to Henry's law, and the chemical equilibria for the dissolution is,

$H_2O + CO_2$ (g) D H_2CO_3 (aq)

Of course, H_2CO_3 dynamically exchanges H^+, and H_2O with other water molecules, and this weak diprotic acid ionizes in two stages with their acid constants, K_{a1} and K_{a2}.

$H_2O + CO_2$ (aq) D H^+ (aq) + HCO_3^- (aq), $K_{a1} = 4.30 \times 10^{-7}$ (at 298 K)

HCO_3^- (aq) D H^+ (aq) + CO_3^{2-} (aq), $K_{a2} = 5.61 \times 10^{-11}$
Constants K_{a1} and K_{a2} increase with temperature. At 298 K, the pH of a solution containing 0.1 mol L^{-1} H_2CO_3 is 3.7, acidophilic organisms may grow, but most pathogenic organisms are neutrophiles and they cease growing. Soft drinks contain other acids – citric, malic, phosphoric, ascorbic acids etc; they lower the pH further.

Ammonia and many nitrogen-containing compounds are weak bases. The ionization equilibrium of NH_3 in water and the base dissociation constant K_b are,
$NH_3 + H_2O$ D NH_4^+ (aq) + OH^- $K_b = 1.70 \times 10^{-5}$ at 298 K.
Other weak bases react with H_2O similarly.

The ionization or dissociation constants of inorganic and organic acids and bases are extensive, and they have been tabulated in various books (39, 40, 41).

Amino acids and proteins contain acidic and basic groups. At some specific pH called the **isoelectric point**, they carry no charge, but exist as zwitterions. For example, the isoelectric point for glycine is pH = 6.00 and it exists as the zwitterion $H_2C(NH_3^+)COO^-$.

C. Oxidation-Reduction Reactions in Water

Oxidation of hydrogen by oxygen not only produces water, but also releases energy. At the standard conditions, the electrochemical half reaction equations are:

$H_2 = 2\ H^+ + 2\ e^-\ E^o = 0.000$ V (defined)

$O_2 + 4\ H^+ + 4\ e^- = 2\ H_2O\ E^o = 1.229$ V

The cell reaction and the cell potential at the standard condition for it are:

$2\ H_2 + O_2 = 2\ H_2O\ DE^o = 1.229$ V

Proper setups for harvesting this energy are the goals of hydrogen-fuel-cell technology. The cell potential DE for non-standard conditions depends on pH and temperature. Its value is related to the energy released in the reaction. A plot of DE versus pH yields a Pourbaix diagram, which is useful to evaluate the stability of various species in water. Water can be a reducing or oxidizing reagent, because it offers protons or electrons. Applying a voltage to pass electrons through a chemical cell decomposes water by electrolysis.

Waters containing dissolved oxygen cause additional reactions, for example:

$2\ H_2O + 2\ e^- = H_2 + 2\ OH^-\ E^o = -0.828$ V

$O_2 + 2\ H^+ + 2\ e^- = H_2O_2\ E^o = 0.682$ V

$O_2 + H_2O + 2\ e^- = HO_2^- + OH^-\ E^o = 0.076$ V

$O_2 + 2\ H_2O + 2\ e^- = 4\ OH^-\ E^o = 0.401$ V

At the proper conditions, a suitable chemical reaction driven by the potential takes place.

Oxidation-reduction reactions involving water usually are due to proton or electron transfer. These oxidation-reduction reactions occur for the growth, production, manufacture, digestion, and metabolism of food.

Water participates in oxidation-reduction reactions in many steps of photosynthesis, resulting in the fixation of CO_2 into biomolecules, releasing oxygen atoms of water as O_2. Engineering a new generation of plants with greater photosynthetic capacity facing lack of waters challenges geneticists (42) and botanists. We now understand photosynthesis to great details, from the studies by many scientists. Photosynthetic reactions are related to food production, but they are so complex that we can only mention them (43).

The oxidation-reduction reactions of water cause corrosion on metal surfaces. Not only deterioration of facilities is very costly for the food industry, corrosion of pipes results in having toxic metal ions Cu^{2+} and Pb^{2+} in drinking water. The concern of lead ions in drinking water led the Environmental Protection Agency to ban the use of high-lead solders for water pipes. These reactions are electrochemical processes. Galvanic effects, high acidity, high flow rate, high water temperature, and the presence of suspended solids accelerate corrosion, as do lack of Ca^{2+} and Mg^{2+} ions in purified waters. The formation of scales protects the metal surface. However, balancing the clogging against surface protection of pipes is a complicated problem, requiring scientific testing and engineering techniques for a satisfactory solution.

D. The Hydrogen Bond and Chemical Reactions

Enzymes are mostly large protein molecules, and they are selective and specific catalysts responsible for most of reactions in biological bodies. Folding of the long protein provides specific 3-dimensional selective pockets for their substrates. The pockets not only fix the substrates in position, they also weaken certain bonds to facilitate specific reactions. This is the mechanism by which enzymes select their substrates and facilitate their specific reactions.

Hydrogen bond strength is stronger in nonaqueous media than in aqueous solutions as the charge densities on the donor and acceptor atoms increase (44). Hydrogen bonds between the enzyme and its substrate can be stronger than those in an aqueous environment, thus speeding up the reaction rate even further.

The hydrolysis of peptide linkage is the reaction of a protein with water,

R-C(=O)-NH-R' + H_2O ® R-C(=O)OH + H_2N-R'

This type of reaction can be catalyzed by acids, bases, and enzymes.

VI. Water Activity

Water is a nutrient and a component in food groups: grains, meat, dairy, fruits and vegetables. Furthermore, major nutrients such as carbohydrates, proteins, water-soluble vitamins, and minerals are hydrophilic. Even parts of fat or lipid molecules are hydrophilic, but the alkyl chains of fats and proteins experience the hydrophobic effect in an aqueous environment. (45).

Foodstuffs interact with water by means of polar, hydrogen-bonding and hydrophobic interactions. The results of these interactions change the chemical potential (properties) of water. Foodstuffs dissolve in or absorb water. Thus, water within food may be divided into bound water, affected water, and free water in the order of their interaction strength. The bound water molecules are similar to those in the first hydration sphere of ions, and those close to the first sphere are affected water molecules. Further away from the interface are free water molecules. The structure and properties of the first two types change. Interaction of water with dietary fiber is an example (46). Thus, properties of water in food are different from those of pure water.

Water molecules in both liquid and vapor phases can participate in hydration reactions. At equilibrium in a system with two or more phases, their vapor pressure or chemical potential, m, must be equal. The chemical potential, m, of a solution or water-containing foodstuff must be equal at a given temperature *T*, and

$m = m_w + R\,T \ln (p / p_w)$,

where *R* is the gas constant (8.3145 J mol^{-1} K^{-1}), and *p* is the vapor pressure of the solution or of water in foodstuff, and p_w is the vapor pressure of pure water at the same temperature. The ratio p / p_w is called the **water activity** $a_{w\,(= p / p_w)}$, which is related to the water chemical potential of water in solutions or in the foodstuff. For ideal solutions and for most moist foods, a_w is less than unity, $a_w < 1.0$ (31).

Both water activity and relative humidity are fractions of the vapor pressure of pure water. Methods for their measurements are the same. We have mentioned the measurement by changes in capacitance earlier. Water contents have a sigmoidal relationship with water activities, $a_w = 1.0$ for infinitely dilute solutions, $a_w > 0.7$ for dilute solutions and moist foods and $a_w < 0.6$ for dry foods. Of course, the precise relationship depends on the material in question. In general, if the water vapor of the atmosphere surrounding the food is greater than the activity of the food, water is absorbed. Otherwise, dehydration takes place. The water activity reflects the combined effects of water-solute, water-surface, capillary, hydrophilic and hydrophobic interactions. The water activity of a foodstuff is a vital parameter, because it affects its texture, taste, safety, shelf life, and appearance.

Furthermore, controlling water activity, rather than water content is important. When $a_w < 0.9$, most molds are inhibited. Growth of yeasts and bacteria also depends on a_w. Microorganisms cease growing if $a_w < 0.6$.

VII Water Potential

Similar to water activity in food, **water potential** is a term used in plant, soil, and crop sciences. Water potential, represented by Y (psi) or Yw, is a measure of the free energy of water in a system: soil, material, seeds, plants, roots, leaves, or an organism. Water potential is the difference between the chemical potential of pure water and water in the system at the same temperature. Pure water has the highest free energy: Y = 0 for pure water by convention, and Y < 0 for solutions. Water diffuses from high potential to low potential. Physiological processes decrease as the water potential decreases.

In general, water potential, Yw ,is a combined effect of osmotic (Ys), matrix (interface and water binding Ym), turgor (Yt) pressures, and gravity (Yg).

Yw = Ys + Ym + Yt + Yg

Osmotic pressure, Ys, is always present due to solutes in the fluids. The metric pressure, Ym, is related to bound-, affected- and free-waters in the system. The outwardly directed pressure extended by the swelling protoplast against the wall is called turgor pressure, Yt. Usually, this term is insignificant until the cell is full, and at such point, Yt increases rapidly and stops when Yw = Yt. Otherwise, the cell ruptures. The mechanical rigidity of succulent plant parts, the opening of stomata and the blossom are usually the results of turgor pressure. In systems such as tall plants and soil science, pressure due to the gravitational pull of water, Yg, is also included in the water potential.

For example, the water potential of potato tissues can be measured by incubating them in a serious of solutions of known osmotic pressures. The potato will neither lose nor gain water if the osmotic pressure of the solution equals the water potential of potato tissues. The osmotic pressure ($p = -mRT$) may be evaluated from a known concentration m, using the equation given earlier. Instead of energy units, water potential is often expressed in units of pressure (megapascal, MPa), which is derived by dividing the energy by the molar volume (0.018 L mol^{-1} for H_2O) of water (47).

There are many other methods for water potential measurements depending on the system: soil, leaf, stem, organism, etc. The soil water potential is related to the water available for the plants growing on the soil. Water potential of a plant or leaf indicates its health or state with respect to water. Thus, water potential is a better indicator for plant, agriculture, irrigation, and environmental managements than water content. Water moves through plants because

$Y_{water} = 0 > Y_{soil} > Y_{root} > Y_{stem} > Y_{leaf}$

Thus, the concept of water potential and water activity are very useful in growth, manufacture, handling, storage, and management of food.

VIII. Living Organisms in Water

The closer we look, the more we see. Living organisms on Earth are so complicated that their classification and phylogeny are still being studied and revised. New relationships are proposed to modify the five kingdoms proposed by Robert Whittaker in 1969. Nevertheless, most of the earliest unicellular living organisms in the Monera, and Protista kingdoms are still living in water. Both the numbers of species and individuals are staggering. For example, photosynthesis by algae in oceans consumes more CO_2 than that by all plants on land. Algae were probably present on Earth before other organisms. Many phyla (divisions) of fungi, plantae, and animalia kingdoms also make water their homes. Both numbers and species of organisms living in water are probably more than those on land. The subject on living organisms in water is fascinating, but we can only mention some fundamentals about their relationships to water here. Certainly, every aspect of living organisms in water is related to food, because Homo sapiens is part of the food chain, if not at the top of it.

All life requires energy or food. Some living organisms receive their energy from the sun whereas others get their energy from chemical reactions in the aquatic media. Chemical reactions are vital during their lives. For example, some bacteria derive energy by catalyzing the oxidation of iron sulfide, FeS_2, to produce iron ions $Fe(H_2O)_6^{2+}$ and elemental sulfur. Water is the oxidant, which in turn reduces oxygen (48). Chemical reactions provide energy for bacteria to sustain their lives and to reproduce. Factors affecting life in water are minerals, solubility of the mineral, electrochemical potentials of the material, acidity (pH), sunlight, dissolved oxygen level, presence of ions, chemical equilibria, etc. Properties of water influence life in general, and in the aquatic system in particular. As the population grows, aquaculture probably will be seen as a more efficient way of supplying protein for the ever-increasing population.

Regarding drinking water, we are concerned with aquatic organisms invisible to the naked eye. Pathogenic organisms present in drinking water cause intestinal infections, dysentery, hepatitis, typhoid fever, cholera, and other illnesses. Pathogens are usually present in waters contaminated with human and animal wastes that enter the water system via discharge, run offs, flood, and accidents at sewage treatment facilities. Insects, rodents, and animals can also bring bacteria to the water system (49, 50). Testing for all pathogenic organisms is impossible, but some organisms have common living conditions to some pathogenic bacteria. Thus, water testing can use these harmless bacteria as indicators for drinking water safety.

IX. Water Resource, Supply, Treatment, and Usage

About 70% of the Earth surface is covered with water, but only about 2% is covered by fresh water. Ocean waters are salty, and only the small percentages that is fresh water resources (lakes, rivers, and underground). Fresh water is needed for drinking, food, farming, washing, and manufacturing.

When salty water freezes, the ice so formed contains very little salt, if any. Thus, nearly all ice, including the massive ice at the polar cap, is fresh water. In fact, the ice cap in the Antarctic contains a lot of fresh-water ice, but that cannot be considered a water resource.

Hydrologists, environmentalists, and scientists, engineers, sociologists, economists, and politicians are all concerned with problems associated with water resources. Solutions to these problems require experts and social consensus.

X. Subcritical and Supercritical Waters

Waters at temperatures between the normal boiling and critical points (0 to 373.98° C) are called **subcritical waters**, whereas the phase above the critical point is **supercritical water**. In the 17th century, Denis Papin (a physicist) generated high-pressure steam using a closed boiler, and thereafter pressure canners were used to preserve food. Pressure cookers were popular during the 20th century. Analytical chemists have used subcritical waters to extract chemicals from solids for analysis since 1994 (51).

Water vapor pressures up to its critical point are given in Table 8, but data on polarity, dielectric constant, surface tension, density and viscosity above 100ºC are scarce. In general, these properties decrease as the temperature increases. In fact, some drop dramatically for supercritical water. On the other hand, some of them increase with pressure. Thus, properties of sub- and super-critical waters can be manipulated by adjusting temperature and pressure to attain desirable properties.

As the polarity and dielectric constant decrease, water becomes an excellent solvent for non-polar substances such as those for flavor and fragrance. However, foodstuffs may degrade at high temperatures. Applications of sub- and super-critical water are relatively recent events, but applications of supercritical CO_2 (critical temperature 32ºC) for chemical analyses started in the 1980s, and investigations of supercritical water followed. However, research and development have been intensified in recent years (52). Scientists and engineers explore the usage of supercritical water for waste treatment, polymer degradation, pharmaceutical manufacturing, chromatographic analysis, nuclear reactor cooling, etc. Significant advances have also been made in material processing, ranging from fine particle manufacture to the creation of porous materials.

Water has been called a green solvent compared to the polluting organic solvents. Sub- and super-critical waters have been explored as replacement of organic solvents in many applications including the food industry (53). However, supercritical water is very reactive, and it is corrosive for stainless steels that are inert to ordinary water. Yet, the application of sub- and super-critical waters is a wide-open field.

XI Postscript

Water, ice, and vapor are collections of H_2O molecules, whose characteristics determine the properties of all phases of water. Together and in concert water molecules shape the landscape, nurture lives, fascinate poets, and captivate scientists. Human efforts in understanding water have accumulated a wealth of science applicable in almost all disciplines, while some people take it for granted.

Water molecules are everywhere, including outer space. They not only intertwine with our history and lives, they are parts of us. How blessed we are to be able to associate and correlate the phenomena we see or experience to the science of water.

An article has a beginning and an end, but in the science of water, no one has the last word, as research and exploration, including its presence in outer space, on water continue (54). Writing this article induced my fascination on this subject, and for this reason, I am grateful to Professors Wai-kit Nip, Lewis Brubacher, and Peter F. Bernath for their helpful discussions and encouragement.

Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

This page titled 2.2.2.5: Fundamental Characteristics of Water is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Chung (Peter) Chieh.

2.2.2.6: Natural Water

Discussion Questions

- What is hard water?
- What are the differences between temporary and permanent hard water?
- How can hard water be converted to soft water?
- How to produce deionized water?

Water is the most important resource. Without water life is not possible. From a chemical point of view, water, H_2O, is a pure compound, but in reality, you seldom drink, see, touch or use pure water. Water from various sources contains dissolved gases, minerals, organic and inorganic substances.

The Hydrosphere

The total water system surrounding the planet Earth is called the **hydrosphere**. It includes freshwater systems, oceans, atmosphere vapour, and biological waters. The Arctic, Atlantic, Indian, and Pacific oceans cover 71% of the Earth surface, and contain 97% of all water. Less than 1% is fresh water, and 2-3 % is ice caps and glaciers. The Antarctic Ice Sheet is almost the size of North America continent. These waters dominate our weather and climate, directly and indirectly affecting our daily lives. They cover 3.35×10^8 km^2. The four oceans have a total volume of 1.35×10^9 km^3.

- The sunlight dims by 1/10 for every 75 m in the ocean, and humans barely see light below 500 m. The temperature of almost all of the deep ocean is 4°C (277 K).
- The average ocean depth is 4 km, and the deepest point at the Mariana Trench is 10,912 m (35,802 ft), which compares to the height of 8.8 km for Mount Everest.

Hydrospheric processes are steps by which water cycles on the planet Earth. These processes include sublimation of ice, evaporation of liquid, transportation of moisture by air, rain, snow, river, lake, and ocean currents. All these processes are related to the physical and chemical properties of water, and many government agencies are set up to study and record phenomena related to them. The study of these processes is called **hydrology**

Among the planets, Earth is the only one in which there are solid, liquid and gaseous waters. These conditions are just right for life, for which water is a vital part. Water is the most abundant substance in the biosphere of Earth. Groundwater is an important part of the water system. When vapor is cooled, clouds and rain develop. Some of the rain percolate through the soil and into the underlying rocks. The water in the rocks is **groundwater**, which moves slowly.

A body of rock, which contains appreciable quantities of water, is called an **aquifier**. Below the **water table**, the aquifier is filled (or saturated) with water. Above the water table is the unsaturated zone. Some regions have two or more water tables. These zones are usually separated by water-impermeable material such as boulder and clay. Groundwater can be brought to the surface by drilling below the water table, and pumped out. The amount of water that can be pumped out depends on the structure of the aquifer. Little water is stored in tight granite layers, but large quantities of water are stored in limestone aquifier layers. In some areas, there are under ground rivers.

Table 2.2.2.6.1: Ions in sea water

Species	Cl^-	Na^+	SO_4^-	Mg^{2+}	Ca^{2+}	K^+	HCO_3^-	Br^-	Sr^{2+}	BO_4^{3-}	F^-	H_4SiO_4	H^+
mg/Kg	10,760	2,710	2,710	1,290	411	399	142	67	8	4.5	1.3	0.5-10	$10^{-8.35}$

Common Ions Present in Natural Water

Hydrology is also the study of how solids and solute interact in, and with, water. In this link, the compositions of seawater, composition of the atmosphere, compositions of rain and snow, and compositions of river waters and lake waters are given in details. Table 2.2.2.6.1 list the major ions present in seawater. The composition does vary, depending on region, depth, latitude, and water temperature. Waters at the river mouths contain less salt. If the ions are utilized by living organism, its contents vary according to the populations of organisms.

Dust particles and ions present in the air are nucleation center of water drops. Thus, waters from rain and snow also contain such ions: Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, NH_4^{+}. These cations are balanced by anions, HCO_3^{-}, SO_4^{-}, NO_2^{-}, Cl^{-}, and NO_3^{-}. The pH of rain is between 5.5 and 5.6. Rain and snow waters eventually become river or lake waters. When the rain or snow waters fall, they interact with vegetation, top soil, bed rock, river bed and lake bed, dissolving whatever is soluble. Bacteria, algae, and water insects also thrive. Solubilities of inorganic salts are governed by the kinetics and equilibria of dissolution. The most common ions in lake and river waters are the same as those present in rainwater, but at higher concentrations. The pH of these waters depends on the river bed and lake bed. Natural waters contain dissolved minerals. Waters containing Ca^{2+} and Mg^{2+} ions are usually called **hard water**.

Hard Water

Minerals usually dissolve in natural water bodies such as lakes, rivers, springs, and underground waterways (ground waters). Calcium carbonate, $CaCO_3$, is one of the most common inorganic compounds in the Earth crust. It is the ingredient for both calcite and aragonite. These two minerals have different crystal structures and appearance. This photograph shows crystals of typical Calcite.

Calcium-carbonate minerals dissolve in water, with a solubility product as shown below.

$$CaCO_3 \rightleftharpoons Ca^{2+} + CO_3^{2-} \quad K_{sp} = 5 \times 10^{-9} \tag{2.2.2.6.1}$$

From the solubility product, we can (see example 1) evaluate the molar solubility to be 7.1×10^{-5} M or 7.1 mg/L (7.1 ppm of $CaCO_3$ in water). The solubility increases as the pH decrease (increase acidity). This is compounded when the water is saturated with carbon dioxide, CO_2. Saturated CO_2 solution contains carbonic acid, which help the dissolution due to the reaction:

$$H_2O + CO_2 \rightleftharpoons H_2CO_3 \tag{2.2.2.6.2}$$

$$CaCO_3 + H_2CO_3 \rightleftharpoons Ca^{2+} + 2HCO_3^{-} \tag{2.2.2.6.3}$$

Because of these reactions, some natural waters contain more than 300 ppm calcium carbonates or its equivalents.

The carbon dioxide in natural water creates an interesting phenomenon. Rainwater is saturated with CO_2, and it dissolves limestones. When CO_2 is lost due to temperature changes or escaping from water drops, the reverse reaction takes place. The solid formed, however, may be a less stable phase called aragonite, which has the same chemical formula as, but a different crystal structure than that of calcite.

The rain dissolves calcium carbonate by the two reactions shown above. The water carries the ions with it, sips through the crack of the rocks. When it reached the ceiling of a cave, the drop dangles there for a long time before fallen. During this time, the carbon dioxide escapes and the pH of the water increases. Calcium carbonate crystals begin to appear. Calcite, aragonite, stalactite, and stalagmite are four common solids found in the formation of caves.

Natural waters contain metal ions. Water containing calcium, magnesium and their counter anions are called **hard waters**. Hard waters need to be treated for the following applications.

- Heat transfer carrier in boilers and in cooling systems
- Solvents and reagents in industrial chemical applications
- Domestic water for washing and cleaning

Temporary vs. Permanent Hard Water

Due to the reversibility of the reaction,

$$CaCO_{3(s)} + H_2CO_3 \rightleftharpoons Ca^{2+} + 2HCO_3^{-} \tag{2.2.2.6.4}$$

water containing Ca^{2+}, Mg^{2+} and CO_3^{2-} ions is called **temporary hard water**, because the hardness can be removed by boiling. Boiling drives the reverse reaction, causing deposit in pipes and scales in boilers. The deposits lower the efficiency of heat transfer in boilers, and diminish flow rates of water in pipes. Thus, temporary hard water has to be softened before it enters the boiler, hot-water tank, or a cooling system. The amount of metal ions that can be removed by boiling is called **temporary hardness**

After boiling, metal ions remain due to presence of chloride ions, sulfate ions, nitrate ions, and a rather high solubility of $MgCO_3$. Amount of metal ions that can not be removed by boiling is called **permanent hardness. Total hardness** is the sum of temporary

hardness and permanent hardness. Hardness is often expressed as equivalence of amount of calcium ions in the solution. Thus, water conditioning is an important topic. The value of water treatment market has been estimated to be worth $30 billion.

Lime-soda Softening

Lime-soda softening is the removal of temporary hardness by adding a calculated amount of hydrated lime, $Ca(OH)_2$:

$$Ca^{2+} + 2HCO_3^- + Ca(OH)_{2(s)} \rightarrow 2CaCO_{3(s)} + 2H_2O \tag{2.2.2.6.5}$$

Adding more lime causes the pH of water to increase, and as a result, magnesium ions are removed by the reaction:

$$Mg^{2+} + Ca(OH)_{2(s)} \rightarrow Mg(OH)_{2(s)} + Ca^{2+} \tag{2.2.2.6.6}$$

The extra calcium ions can be removed by the addition of sodium carbonate.

$$Na_2CO_3 \rightarrow 2Na^+ + CO_3^{2-} \tag{2.2.2.6.7}$$

$$Ca^{2+} + CO_3^{2-} + \rightarrow CaCO_{3(s)} \tag{2.2.2.6.8}$$

In this treatment, the amount of $Ca(OH)_3$ required is equivalent to the temporary hardness plus the magnesium hardness. The amount of sodium carbonate required is equivalent to the permanent hardness. Thus, lime-soda softening is effective if both the temporary and total hardness have been determined. The sodium ion will remain in the water after the treatment. The pH of the water is also rather high depending on the amount of lime and sodium carbonates used.

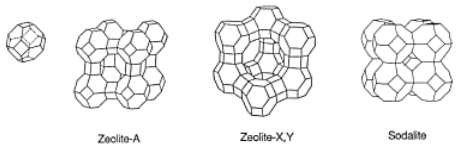
Complexation Treatment

Addition of complexing reagent to form soluble complexes with Ca^{2+} and Mg^{2+} prevents the formation of solid. One of the complexing agents is sodium triphosphate Na_3PO_4, which is marketed as Calgon, etc. The phosphate is the complexing agent. Other complexing agents such as Na_2H_2EDTA can also be used, but the complexing agent $EDTA^{4-}$ forms strong complexes with transition metals. This causes corrosion problem, unless the pipes of the system are made of stainless steel.

Ion Exchange

Today, most water softeners are using zeolites and employing ion exchange technique to soften hard water. Zeolites are a group of hydrated crystalline aluminosilicates found in certain volcanic rocks. The tetrahedrally coordinated aluminum and silicon atoms form AlO_4 and SiO_4 tetrahedral groups. They interconnect to each other sharing oxygen atoms forming cage-type structures as shown on the right. This diagram and the next structural diagram are taken from an introduction to zeolites There are many kinds of zeolites, some newly synthesized.

Whatever kind, the crystal structure of zeolites contains large cages. The cages are connected to each other forming a framework with many cavities and channels. Both positive and negative ions can be trapped in these cavities and channels as shown below.



For each oxygen that is not shared in the AlO_4 and SiO_4 tetrahedral groups, a negative charge is left on the group. These negative charges are balanced by trapping alkali metal and alkaline earth metal ions. When more cations are trapped, hydroxide and chloride ions will remain in the cavities and channels of the zeolites.

To prepare a zeolite for water treatment, they are soaked in concentrated NaCl solution. The cavities trap as many sodium ions as they can accommodate. After the treatment, the zeolite is designated as Na-zeolite. Then the salt solution is drained, and the zeolite is washed with water to eliminate the extra salt. When hard water flow through them, calcium and magnesium ions will be trapped by the Na-zeolite. For every Ca^{2+} or Mg^{2+} trapped, two Na^+ ions are released. The treated water contains a rather high concentration of Na^+ ions, but low concentrations of Mg^{2+} and Ca^{2+}. Thus, zeolite ion exchange convert hard water into soft water.

Pure Water by Ion Exchange

In most cases, the resins are polystyrene with functional -SO_3H groups attached to the polymer chain for cation exchange resin, and with functional group -$N(CH_3)_3^+$ attached to the chain for anion exchange resin. To prepare the resin for making *pure or deionized water*, the cation resin is regenerated with HCl so that the groups are really -SO_3H. The anion resin is regenerated with NaOH, so that the functional groups are -$N(CH_3)_3(OH)$. When water containing any metal ion M^+ and anion A^- passes through the ion exchange resins in two stages, the following reactions take place,

$$M^+ + -SO_3H \rightarrow H^+ + -SO_3M \tag{2.2.2.6.9}$$

$$A^- + -N(CH_3)_3(OH) \rightarrow OH^- + -N(CH_3)_3A \tag{2.2.2.6.10}$$

$$H^+ + OH^- \rightleftharpoons H_2O \tag{2.2.2.6.11}$$

Thus, ion exchange provides *pure water* to meet laboratory requirement.

Reverse Osmosis Water Filter System

This method can also be used to prepare water for domestic and laboratory applications. This method has been discussed in Wastewater treatment

Magnetic Water Treatment

The following is a list of companies selling magnetic devices for **magnetic water treatment**. All devices are based on results of some ressearch indicating that when water runs through a magnetic field, the calcium carbonate will precipitate as aragonite rather than the usual calcite. For example, K.J. Kronenberg has published an article in *IEEE Transactions on Magnetics*, (Vol. Mag-21, No. 5, September 1985, pages 2059-2061). and stated the following:

> *The crystallization mode of the water's mineral content was found to change from a dendritic, substrate-bound solidification habit to the form of separate disc-shaped crystals after the water had moved through a number of magnetic fields. The former scarcity of crystallization nucleii in the water had been turned into an abundance of nucleation centers in the water. The reduction of the number of the substrate-bound crystals has been used as a quantitative measure of the magnetic effect.*

Many companies have made various devices for magnetic conditioning of water, and they claim that their devices will clean up the pipes and boilers at little or no cost. I have yet to test one of these devices for its claim, but my preliminary tests shows that permanent magnet has little effect on the calcium carbonate deposit of temporary hard water. The cleaning effect they have claimed is probably much overstated.

Example 2.2.2.6.1

From the solubility product shown for the dissolution of calcium carbonate,

$$CaCO_3 \rightleftharpoons Ca^{2+} + CO_3^{2-} \quad K_{sp} = 5 \times 10^{-9} \tag{2.2.2.6.12}$$

Evaluate the molar solubility of Ca^{2+} in saturated solution.

Solution

From the definition of solubility product, we have

$$[Ca^{2+}][CO_3^{2-}] = 5 \times 10^{-9} \tag{2.2.2.6.13}$$

Thus,

$$[Ca^{2+}] = [CO_3^{2-}] = 7.1 \times 10^{-5}\ M \tag{2.2.2.6.14}$$

The concentration of 7.1×10^{-5} M is equivalent to 7.1 mg/L (7.1 ppm of $CaCO_3$ in water).

DISCUSSION

There may be other ions present in the system and other equilibria conditions in addition to the equilibrium mentioned here. Problems are more complexed in the real world.

Exercise 2.2.2.6.1

Boiling of 1.0 L of water produced 10 mg of $CaCO_3$ solid. What is the temporary hardness of the water?

Answer

10 ppm

Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

This page titled 2.2.2.6: Natural Water is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Chung (Peter) Chieh.

2.2.2.7: Water Biology

Discussion Questions

- Is water the source of life, why?
- What is life?
- Is life a symbiotic system of many lives?
- What properties of water make water the prime ingredient of life?
- What are the essential substances of life?
- Other than material, what else is essential for life?
- What is the role of water in supporting life?
- Is there a life form involving no water?
- Is there a method to determine the total water content?

Not finished yet, but the major ideas are here now!

Water Biology



Chemical and physical properties of water discussed in other pages are essential considerations for water biology. Natural water also contains biological matters as well as living creatures. In the discusssion of biology and water, you must feel pleasant, because they make each other grow more lovely as this picture from water lily cottage. The water pages of National Wildlife Federation, offers some interesting reading too.

All living things have a cycle of life. A cycle involves all or some of these process: birth, growth, mature, reproduction, matamorphosis, and death. There are millions of living organisms on Earth, ranging from single-cell amebas, bacteria, to the complicated homosapiens. There are also viruses which are fragment of DNA or RNA that depend on host cells for their reproduction. They are not cells.

Living things usually have cells that isolate their systems so that the cells contain unique materials to sustain the lives of cells. Cells regulate their contents (homeostatic), and carry out their metabolsims. They divide or making copies of themselves. Many reproduction process involve two individuals and future populations are subject to a greater diversity. Mutation is a fact of life, and many adopt to their changing environment.

How life started? Let the research and debate continue by not giving any clusive statements here. A physical geography course suggests that the marine invertebrates began their life 600 million years ago, and they are followed by fish, land plants, amphibians, reptiles, mommals, and then flowering plants, in this order. All these started more than a hundred million years ago, and hominid (primate) line began its evolution 20 to 15 million years ago.

There are strong evidences that life on earth appeared in a body of water. Only the planet Earth has three states of water, and it offers a suitable environment for life to began, among all nine solar planets. Since all life forms involve water. Water is seen as the source, matrix, and mother of life. Water is important, because water is required for life, and some people even consider water as life blood.

Since water supports life, living organisms also modify their environment, changing the nature of the water in which they live. Biology of water pollution, lists the syllabus on a course including a laboratory section. Water and biology interweave into an entangled maze waiting for explorers and curious minds.

Water dissolves or emulsifies other life-supporting substances and transport them to intercellular and intracellular fluids. It is also a medium in which reactions take place. Reactions provide energy (non-matter) for living. Energy causes changes, and manifestation of changes is at least related to, if not the whole, life. An organized and systematized set of reactions is essential in each life.

Balancing Water in Bio-systems

Many living organisms live their lives entirely in water as shown here in this photo from a job center talking about work in marine biology. Aquatic living organisms extract neutrients from water, yet maintaining a balance of electrolyte and nurrishment concentration in their cells. For living things not living in water, they extract water from their environment by whatever mechanism they can. Cells in their body are surrounded by body fluid, and all cells maintain constant concentrations of electrolytes, neutrints, and metabolites. The process of maintaining constant concentrations is called homeostasis. Certainly, some active transport mechanisms are involved in this balance.

The rooting of every type of plants is unique. Generally speaking, plants having extensive roots are able to extract water under harsh conditions. On the other hand, some plants such as cactus, jade and juniper have little roots, but their leaves have a layer of wax that prevents water from evaporation. Water conserving plants tolerate draught, and they survive under harsh conditions. The picture shown here is a jade plant from the above link.

Lately, some pumpkin growers harvested squash weighing almost 500 kg. At the peak of the growing season, the squash grows almost 0.5 kg a day. That is equivalent to 25 moles of water collected by the roots, discounting the water evaporated through the leaves. The growth is particularly good during a hot and wet day, but during a hot sunny after noon, the temperature of the leaves and fruits get very hot.

Essential Electrolytes for Life Support

In addition to water, many inorganic substances or minerals are essential to life. These substances ionize in water to form ions and their solutions conduct electricity. Therefore, they are called **electrolytes**. Because most of these substances are already dissoved in natural water, we list ions instead of the mineral they come from.

When ions dissove, they form complexes with water molecules. For most metals, the first sphere of coordination usually involve 6 water molecules. For example, when sodium chloride dissolve, we have

$$\mathrm{NaCl} + 12\,\mathrm{H_2O} \rightleftharpoons \mathrm{Na(H_2O)_6}^{+} + \mathrm{Cl(H_2O)_6}^{-} \tag{2.2.2.7.1}$$

$$\mathrm{FeCl_2} + 18\,\mathrm{H_2O} \rightleftharpoons \mathrm{Fe(H_2O)_{62}}^{+} + 2\,\mathrm{Cl(H_2O)_6}^{-} \tag{2.2.2.7.2}$$

Formation of complexes are due to the high dipole moment of water, and the dissolution can be attribute to the high dielectric constant (80). However, in most publications, we ignore the water molecules in the complexes, and simply consider them as ions.

In the following, we describe some essential ions or salts as electrolytes.

- Sodium chloride, Na^+ and Cl^-
  NaCl is readily dissolved and absorbed in extracellular fluid. The two ions help to balance water, acid/base, osmotic pressure, carbon dioxide transport, and excreted in human urine and sweat. Lack of sodium chloride shows symptoms of dehydration.
- Potassium, K^+
  Good sources of potassium ions are vegetables, fruits, grains, meat, milk, and legumes. It is readily absorbed, and actively transported into the intracellular fluid. Its function is similar to that of sodium ions, but cells prefer potassium ions over sodium. Lack of potassium leads to cardiac arrest.
- Calcium, Ca^{2+}
  Divalent calcium ions are usually poorly absorbed by human, but the are essential for the bones, teeth, blood clotting. Lack of calcium hinders growth and osteoporosis in old age.
- Phosphates, PO_4^{3-}
  Calcium phosphate is essential for bones, teeth, etc. However, phosphates are also responsible for many life reactions. ATP, NAD, FAD etc are metabolic intermediates, and they involve phosphate. Phospholipids and phosphoproteins are some other phosphate containing species.
- Magnesium, Mg^{2+}
  Magnesium ions are essential in chlorophyll. These ions are absorbed readily, and compete with calcium at times. Magnesium and calcium ions are present in hard water, and this link alerts the lack of magnesium leading to cardiovascular disease.
- Ferrous or ferric ions, Fe^{2+} or Fe^{3+}
  Usually known as iron, but iron are present either as divalent or as trivalent ions. Iron is absorbed according to body need; aided

by HCl, ascorbic acid (vitamin C), and regulated by apoferritin. In mammals, iron is stored in liver as ferritin and hemosiderin. Iron deficiency leads to anemia. Good food sources of iron are liver, meats, egg yolk, green vegetables, and whole grains.

- Zinc ions, Zn^{2+}
  Zinc ions are important ingredients for many enzymes. They are present in insulin, carbonic anhydrase, carboxypeptidase, lactic dehydrogenase, alcohol dehydrogenase, alkaline phosphatase etc. Like iron, zinc deficiency leads to anemia and poor growth.
- Copper ions, Cu^{2+}
  Copper ions help iron utilization, and this metal is present in may enzymes.
- Cobalt ions, Co^{2+}
  Cobalt ions are centers of vitamin B12, and deficiency of which leads to anemia.
- Iondine ions, I^{-}
  Iondine is a constituent of thyroxin, which regulates cellular oxidation.
- Fluoride ions, F^{-}
  Fluoridation of drinking water is often a controversal issue. Childen's teeth are less seceptible to decay. Once they began to bruch their teeth, the fluoride in tooth paste is sufficient.

The eletrolytes listed above are present in significant amount in water, or fluids of organisms. There are some metals present in very minute quantities in biological systems, and these are not listed above.

Metal ions also interact with proteins. An **enzyme** is usually a very large protein molecule, and it folds into a kidney shape enclosing one or more metal ions forming a complex. The metal is usually responsible for the enzyme activity. Cobalt, copper, iron, molebdenium, nickel, and zinc have groups of enzymes each, and further discussion can be found in The Prosthetic groups and Metal Ions in Protein Active Sites Database (PROMISE). A general discussion is called **bioinorganic chemistry** and this site has an extensive General references on Bioinorganic Chemistry.

Balancing Electrolytes

Ions	Extracellular	Intracellular	Interstitial
Na^{+}	140	10	150
K^{+}	5	150	4
Ca^{2+}	10	4	6
Mg^{2+}	6	80	4
Total	161	244	164
Cl^{-}	103	2	120
$HCO3^{-}$	30	10	30
$HPO4^{2-}$	4	177	4
$SO4^{2-}$	2	10	2
Organic acid	6	5	6
Protein	16	40	2
Total	161	244	164

Electrolyte balance are maintained by **passive transport or diffusion and slective active transport mechanisms**. Diffusion process tends to make the concentration all the same throughout the entire fluid, but active or selective transport moves ions to special compartment. For example, the active transport of sodium and potasium by an enzyme called **sodium-potassium ATPase** is usually known as *sodium-potassium pump*. This process pumps potassium ions inside a cell while removing sodium ions from the cells. Thus, a high concentration of potassium is maintained inside cells. Energy is required in active transport, and cellular metabolism provides the energy and the necessary molecular motions to facilitate the process.

Hormons are produced by special cells, and they are responsible for the communication between various part of the body. Some complicate harmon actions regulates the rate of transport and balance the ion concentrations depending on the portion of the tissue and the need. This is generally called the **hormonal effects** following the suggestion of *human biochemistry*.

Gibbs-Donnan effect considers the equilibrium in compartments that are separated by memberances or cell walls. There will be no net change when the products of concentrations of say $[Na^+]_1$, $[Cl^-]_1$ are the same for compartments 1 and 2.

$$[Na+]1\ [Cl-]1 = [Na+]2\ [Cl-]2$$

the subscript 1 and 2 refer to the two compartments. When no other components are present, we have

$$[Na+]1 = [Cl-]1 = [Na+]2 = [Cl-]2$$

But if compartment 2 has a sodium salt with other anions, this salt ionize to give Na^+ too. The above condition will not be maintained, in this case. In other words, thermodynamic will be a force to adjust the concentrations.

In general, the cations should be balanced with anions. Otherwise, the solution will be charged.

Water in Human Biology

In human, water in the tissue and body fluid is mostly free, but some fraction may be bounded in pockets of hydrophilic compartments. Body fluids have many electrolytes and neutrients dissolve in them.

Intracellular fluid	70%
Interstitial fluid (lymph)	20%
Blood plasma	7%
Intestinal lumen etc.	3%

Human Biochemistry by J.M. Orten and O.W. Neuhaus (1982), 10th Ed. suggests that about 70% of human body weight is water, most found in three major compartments: 70% intracellular fluid, 20% interstitial fluid, and 7% blood plasma, and only 3 % in intestinal lumen, cerebrospinal fluid and other compartments.

However, *Human Biochemistry* also suggests that blood makes up about 8% of the total body weight.

Example 1

For a person weighing 50 kg (110 lb), what is the weight of the blood?

Solution

From the distribution given above,

$$\text{Amount of blood} = 50\ \text{kg} * 0.08$$
$$= 4\ \text{kg}.$$

This is a lot of blood, and donating 0.5 L of blood will not affect the normal function of the blood.

Input		Output	
Drinking	400 g	Skin	500 g
Beverage	580	Expired air	350
Preformed water in solid food	720	Urine	1100
Metabolic water	320	Feces	150
Total	2020	Total	2100
Balance	-80 g?		

Water in human comes from ingestion. Aside from drinking water, there is other beverages. Much of the food also contain water. When food is oxidized in the cells, all hydrogen in food converts to water, which is called metabolic water. Water is excreted via urine, feces, skin, and expiration. A typical daily water balance is shown in a table here. Water balance is maintained between cells and fluid, and the output depends on kidney functions and body insensible perspiration (Expired air from the lung is saturated with water vapor, and evaporation from the skin).

Drinking Water

Drinking water affects health. An Excite search using the phrase "drinking water" came up with 57890 documents. Drinking Water Resources gives annotated links to web sites that provide information about the drinking water.

A rather recent book *Chemistry of Water Treatment* by S.D. Faust and O.M. Aly, 2nd Ed. (1998) [TD433 F38 1998] addresses the standards for drinking water in the first Chapter. The standards have changed over the years, as we better understand the science.

Safe drinking water is a suitable combination of minerals and electrolytes. Usually, one should not drink water softened by water softeners. Using distilled water for beverages and cooking may not reach your set goals. Hard water with calcium and magnesium ions is good for drinking.

Usually, a government set up a non-profit organization to provide rules for safe drinking water. This organization has an infrastructure to monitor drinking water systems, and it shall also carry out research to improve the quality of drinking water.

Regarding making rules, reliable tests should be developed to determine the electrolyte we have mentioned, plus others such as lead ions, Pb^{2+}, mercury Hg^{2+}, methylmercury, arsenic, radioactivity, etc. Bacteria tests should be carried out regularly. This organization should also have a communication channel to release relavant message.

The Environmental Protection Agency of the U.S. gives a list of comtaminants. The list has suggested limits, and it divides the contaminants into

- Inorganic substnaces
- Organic substances
- Radioactivities (alpha and beta rays, radium)
- Micro-organisms

Among inorganic substances, limits are given to contents of antimony, arsenic, asbestos, barium, beryllum, cadmium, chromium, copper, mercury, nitrate, nitrite, selenium, and thallium.

More than 50 organic compounds are on the list, and some familar ones are: acrylamide, benzene, carbon tetrachloride, chlorobenzene, 2 4 D, dichlorobenzen, dioxin, polychlorinated biphenyls (PCBs), toluene, and vinyl chloride. Many of these have a zero limit.

In terms of microorganisms, *Giardia lamblia*, and *Legionella*, are checked. Furthermore, viruses, turbidity, total coliforms, and heterotrophic plate should be checked.

The secondary standard lists most electrolytes as shown in this table on the right.

The Secondary Drinking Water Standards are non-enforceable guidelines regulating contaminants that may cause cosmetic effects (such as skin or tooth discoloration) or aesthetic effects (such as taste, odor, or color) in drinking water.

There are many brands of bottled drinking water, they have been very popular only in recent years. Do we know the bottle procedure? Is the industry regulated? Is the water quality reliable? Are all bottled water the same? Is there a brand bottled water for real good health? Do we know what should be in healthy drinking water? There is an opinion expressed in Ontario Clean Water Agency (OCWA). Check it out.

The magnesium web site gave the following news release Oct. 4, 1999. *According to the U.S. National Academy of Sciences (1977) there have been more than 50 studies, in nine countries, that have indicated an inverse relationship between water hardness and mortality from cardiovascular disease. That is, people who drink water that is deficient in magnesium and calcium generally appear more susceptible to this disease. The U.S. National Academy of Sciences has estimated that a nation-wide initiative to add calcium and magnesium to soft water might reduce the annual cardiovascular death rate by 150,000 in the United States.* This is a good summary from the report.

Contaminant	Standard
Aluminum	0.05 to 0.2 mg/L
Chloride	250 mg/L
Color	15 (color units)
Copper	1.0 mg/L
Corrosivity	noncorrosive
Fluoride	2.0 mg/L
Foaming Agents	0.5 mg/L
Iron	0.3 mg/L
Manganese	0.05 mg/L
Odor	3 threshold odor number
pH	6.5-8.5
Silver	0.10 mg/L
Sulfate	250 mg/L
Total Dissolved Solids	500 mg/L
Zinc	5 mg/L

Sports Drinks

Sports talk gives information about Sports Drinks. This is science, art, testing, and myth. However, some fundamentals should be considered.

Our bodies are mostly water, about 70%. The body fluid has many different things dissolved in it, particularly salt. the salinity - varies somewhat with where you take the sample of water to measure. Don't worry about that). I recall the concentration as being 0.5%, but I'm not sure. Doctors call this "normal saline".

Now if you put a human, or other animal, cell in saltwater that is the same concentration as the saltwater inside the cell, the cell pretty much just sits there. If you put it in distilled water, the cell absorbs the water through its cell membrane - called diffusion - until it eventually pops. If you put the cell in concentrated saltwater, the cell looses water. The water diffuses out of the cell through the membrane, leaving a small, shriveled up cell.

What does this have to do with sports drinks? If you give someone distilled water, it seems like they would absorb the water faster because of what I just described. On the other hand, during sweating, you're loosing sodium, potassium and small quantities of other electrolytes. If you're exercising particularly long or hard, you need to replace those electrolytes. Researchers found that adding some salt to water replaced the salt lost through sweating and helped the body to get water to the cells. If you look at a label on a Gatorade or other drink, you'll find that the main electrolyte is simple salt. But if you put too much of the electrolytes in the water, the cells shrivel up just like the way I described above.

I hope that helps you understand what happens. This stuff on how cells swell up or shrivel up is in high school biology books, and maybe even in something you can find in your school's library.

Taste and Order of Drinks

Taste and order are sensations, and thus hard to quantify and systematize. Often, the Weber Fechner law is used. This law expresses the taste or order sensations S as being proportional to the logarithm of the stimulus R, with the proportional constant K,

$$S = K \log R$$

For some common substances, the minimum amounts detected by *expert nose* or *expert taster* gives the sensation as a threshold, below which no taste or order was detected.

However, orderous sensation of water may be reported as threshold order number (TON). If *A* mL of ordorous sample is diluted with *B* mL of order-free water to be "just detectable" to the *expert nose*, the TON is defined as

$$TON = \frac{A + B}{A}$$

Similarly, a flavour threshold number (FTN) can be defined in the same manner.

$$FTN = \frac{A + B}{A}$$

except that *A* and *B* are volumes of samples and taste free water used.

These formulations show a method of defining some quantities that are, otherwise, very diffcult to quantify. There are other methods of reporting orders and taste, and some bottling companies have their standard methods of comparison.

Of course, the source of order and taste comes are organic, inorganic compounds as well as bacteria, algae. For example, mercaptans such as C_2H_2SH and ammonia offers disagreeable smell and taste.

Example 2

A sample of water was tested by 10 expert noses, and only 5 of them can detect an order. Thus, this is "just detected". What is the TON for this sample?

Solution

Since no dilution was used,

$$TON = A / A = 1.$$

DISCUSSION

If equal amount of orderless water is required to dilute it so that the order is "just detected", then the test order number is (1+1)/1 = 2.

Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

This page titled 2.2.2.7: Water Biology is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Chung (Peter) Chieh.

2.2.2.8: Water Chemistry

Water is an unusual compound with unique physical properties. As a result, its the compound of life. Yet, its the most abundant compound in the biosphere of Earth. These properties are related to its electronic structure, bonding, and chemistry. However, due to its affinity for a variety of substances, ordinary water contains other substances. Few of us has used, seen or tested pure water, based on which we discuss its chemistry.

The chemistry of water deals with the fundamental chemical property and information about water. Water chemistry is discussed in the following subtitles.

- Composition of water
- Structure and bonding of water
- Molecular Vibration of water
- Symmetry of water molecules
- Formation of hydrogen bonding in water
- Structure of ice
- Autoionization
- Leveling effect of water and acid-base characters
- Amphiprotic nature
- Reactivity of water towards alkali metals; alkaline earth metals; halogens; hydrides; methane; oxides; and oxygen ions.
- Electrolysis of water

Composition of Water

Water consists of only hydrogen and oxygen. Both elements have natural stable and radioactive isotopes. Due to these isotopes, water molecules of masses roughly 18 ($H_2{}^{16}O$) to 22 ($D_2{}^{18}O$) are expected to form. Isotopes and their abundances of H and O are given below. From these data, we can estimate the relative abundances of all isotopic water molecules.

Abundances (% or halflife) of hydrogen and oxygen isotopes

^1H	^2D	^3T		
99.985%	0.015%	12.33 y		
^{14}O	^{15}O	^{16}O	^{17}O	^{18}O
70.6 s	122 s	99.762%	0.038%	0.200%

The predominant water molecules $H_2{}^{16}O$ have a mass of 18 amu, but molecules with mass 19 and 20 occur significantly. Because the isotopic abundances are not always the same due to their astronomical origin, The isotopic distribution of water molecules depends on its source and age. Its study is linked to other sciences. (*See* Dojlido, J.R. & Best, G.A. (1993) *Chemistry of Water and Water Pollution*, Ellis Harwood for isotopic distribution of water.)

Relative abundance of isotopic water

$H_2{}^{16}O$	$H_2{}^{18}O$	$H_2{}^{17}O$	$HD^{16}O$	$D_2{}^{16}O$	$HT^{16}O$
99.78%	0.20%	0.03%	0.0149%	0.022 ppm	trace
18	20	19	19	20	20 amu

In particular, $D_2{}^{16}O$ is called heavy water, and it is produced by enrichment from natural water. Properties of heavy water are particularly interesting due to its application in nuclear technology.

Structure and Bonding of the Water Molecule

Pure water, H_2O, has a unique molecular structure. The O-H bondlengths are 0.096 nm and the H-O-H angle = 104.5°. This strange geometry can be explained by various methods. From carbon to neon, the numbers of valence electrons increase from 4 to 8. These

elements require 4, 3, 2, 1, and 0 H atoms to share electrons in order to complete the octet requirement. Their **Lewis dot structures** are shown on the right, and note the trend in bondlengths.

There are six valance electrons on the oxygen, and one each from the hydrogen atom in the water molecule. The eight electrons form two H-O bonds, and left two lone pairs. The long pairs and bonds stay away from each other and they extend towards the corners of a tetrahedron. Such an ideal structure should give H-O-H bond angle of 109.5°, but the lone pairs repel each other more than they repel the O-H bonds. Thus, the O-H bonds are pushed closer, making the H-O-H angle less than 109°.

After the introduction of quantum mechanics, the electronic configuration for the valence electron of oxygen are $2s^2\ 2p^4$. Since the energy levels of 2*s* and 2*p* are close, valence electrons have characters of both *s* and *p*. The mixture is called ***sp*^{3} hybridization.** These hybridized orbitals are shown on the right. The structures of CH_4, NH_3, and H_2O can all explained by these hybrid orbitals of the central atoms. The above approach is the valence bond theory, and both the C-H bonds and lone electron pairs are counted as **VSPER pairs** in the Valence-shell Electron-Pair Repulsion (VSEPR) model, according to which, the four groups point to the corners of a tetrahedron.

For triatomic molecules such as water, molecular orbital (MO) approach can also be applied to discuss the bonding. The result however is similar to the valence bond approach, but the MO theory gives the energy levels of the electron for further exploration.

Molecular Vibration of Water

Atoms in a molecule are never at rest, and for each type of molecule, there are some normal vibration modes. For the water molecule, the three normal modes of vibrations are symmetric stretching, bending and asymmetric stretching. The vibrations are quantized, as do any microscopic system, and their quantum numbers are designated as v_1, v_2 and v_3. The observed transition bands of D_2O, H_2O, and HDO are given in the table below.

Transition bands of D2O, H2O, and HDO

Quantum numbers of upper state			Absorption wavenumbers of bands /cm^{-1}		
v_1	v_2	v_3	D_2O	H_2O	HDO
0	1	0	1178	1594	1402
1	0	0	2671	3656	2726
0	0	1	2788	3756	3703
0	1	1	3956	5332	5089

Data from Eisenberg, D. and Kauzmann, W. (1969) *Structure and properties of water*, Oxford University press.

The ideal transition bands are centered in the given wavenumbers. However, these wavenumbers are calculated based on isolated molecules with no interaction with any neighbor. When molecules interact with each other, the energy levels are modified, and the bands shift.

Many more less intense absorption bands extend into the green part of the visible spectrum. The absorption spectrum of water may contribute to the blue color for lake, river and ocean waters.

Symmetry of Water Molecules

The water molecules are rather symmetric in that there are two mirror planes of symmetry, one containing all three atoms and one perpendicular to the plane passing through the bisector of the H-O-H angle. Furthermore, if the molecules are rotated 180° (360°/2) the shape of the molecule is unperturbed. This indicates that the molecules have a 2-fold rotation axis. The three symmetry elements are 2-fold rotation, and two mirror planes. Both mirror planes contain the rotation axis, and this type of symmetry belongs to the **point group** C_{2v}. A **point group** has a definite number of symmetry elements arranged in certain fashion. Molecules can be classified according to their point groups. Molecules of the same point group have similar spectroscopic characters. Other molecules of C_{2v} point group are CH_2=O, CH_2Cl_2, the bent O_3 etc.

Formation of Hydrogen Bonding

Under certain conditions, an atom of hydrogen is attracted by rather strong forces to two atoms instead of only one, so that it may be considered to be acting as a bond between them. This is called ***hydrogen bond***. This statement is from Linus Pauling (1939) in his book *The Nature of the Chemical Bond*. He gave the ion $[F:H:F]^-$ as an example. At that time, the hydrogen bond was recognized as mainly ionic in nature. The energy associated with hydrogen bond is 8 to 40 kJ/mol.

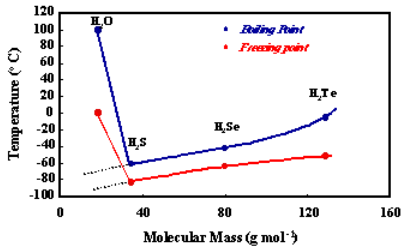
Comparison of melting and boiling points for a few substances

Molecule	Molar mass	m.p.	b.p. /° C
NH_3	17	-77.8	-33.5
H_2O	18	-0	100
H_2S	34	-85.6	-60
H_2Se	81	-60.4	-41.5
H_2Te	128.6	-51	-1.8
CH_3OH	32	?	65
C_2H_5OH	46	?	78
$C_2H_5OC_2H_5$	74	?	34

Normally, the melting point and boiling point of a substance increase with molecular mass. For example the melting points of inert gases are 0.95, 24.48, 83.8, and 116.6 K respectively for He, Ne, Ar, and Kr.

In this table, the melting and boiling points for water are particular high for its small molecular mass. This is usually attributed to the formation of hydrogen bonds. The small electronegative atoms F, O and N are somewhat negatively charged when they are bonded to hydrogen atoms. The negative charges on F, O and N attract the slightly positive hydrogen atoms, forming a strong interaction called hydrogen bond.

A graph showing the melting points and boiling points of group 16 provided by Prof. J. Boucher illustrates the same point.



Boiling and Freezing Points of Group 16 Hydrides

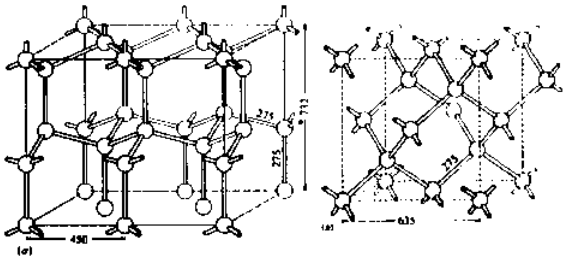
Based on the observed absorption at 3546 and 3691 cm^{-1}, Van Thiel, Becker, and Pinmentel (1957, *J. Chem. Phys.* **27** 386) suggested the formation of water dimer when trapped in a matrix of nitrogen.

Due to hydrogen bonding, water molecules form dimers, trimers, polymers, and clusters. The hydrogen bonds are not necessarily liner.

Structure of Ice

Ice occurs in many places, including the Antarctic. If all the ice melted, the water level of the oceans will rise about 70 m. The structure of ice and the caption are from this link.

The density of ice is dramatically smaller than that of water, due to the regular arrangement of water molecule via hydrogen bonds. In an idealized structure of ice, every hydrogen atom is involved in hydrogen bond. Every oxygen atom is surrounded by four hydrogen bonds.

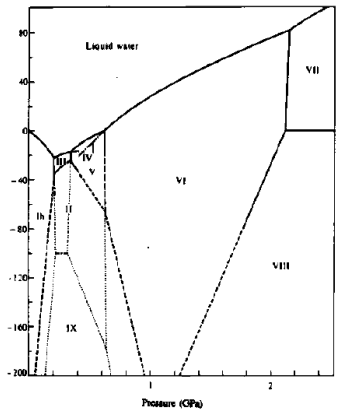


This diagram from caltech.edu, shows the structure of hexagonal ice in (a) and cubic ice in (b). A rod here represents a hydrogen bond. Since the hydrogen bonds are not linear, the real structure is a little more complicated.

The tetrahedral coordination opens up the space between molecules. On each hydrogen bond, shown by a rod joining the oxygen atoms, lies one proton in an asymmetric position (not shown). Bond lengths, 275 pm, are indicated. Ordinary ice is hexagonal. and the hexagonal *c* axis is labelled 732 pm, and one of the hexagonal *a* axes is labelled 450 pm. If water vapor condenses on very cold substrate at 143-193 K (-130 to -80ºC) a cubic phase is formed. In (b) the cubic unit cell is outlined with dashed lines; dimensions are in pm determined at 110 K.

These diagrams can also be used to represent the two forms of diamond, and in this case, the rods joining the atoms represent C-C bonds. Each C-C bondlength is 154 pm. Silicon and germanium crystals have the same structure, but their bondlengths are longer. The two diamond types of structure are related to the packing of spheres. The hexagonal type has the ABABAB... sequence, whereas the cubic type has the ABCABC... sequence. In both cases, half of the tetrahedral sites are occupied by tetrahedrally bonded carbon atoms. Hexagonal diamonds have been observed in meteorites.

The four hydrogen bonds around an oxygen atom form a tetrahedron in a fashion found in the two types of diamonds. Thus, ice, diamond, and close packing of spheres are somewhat topologically related.



A phase diagram of water shows 9 different solid phases (ices). Ice Ih is the ordinary ice. In addition to ice Ic from vapor deposition, conditions for nine phases are shown. Aside from ice I, other phases are formed and observed under high pressure generated by machines built by scientists. So far, ten different forms of ice have been observed, and some ice forms exist at very high pressure. The pressure deep under the polar (Antarctic) ice cap is very high, but we are not able to make any direct observation or study.

There is a report of the 11th ice, and the ice phase diagram and drawings of ice structures given here is extremely interesting.

The Autoionization of Water

The Autoionization of Water in the formation of ions according to

$$HOH(l) + HOH(l) = H3O+ + OH-$$

This is an equilibrium process and is characterised by an equilibrium constant, K'_w:

$$K'w = \frac{[H3O+]\,[OH-]}{[H2O]}$$

Since [H_2O] = 1000/18 = 55.56 M, and remains rather constant under any circumstance, we usually write

Kw = [H3O+] [OH-]
= 10-14 (or 1e-14)
pKw = -log Kw (defined)
= 14 (at 298 K)

t°C	K_w
20	1.14e-15
25	1.00e-14
35	2.09e-14
40	2.92e-14
50	5.47e-14

For neutral water, [H_3O^+] = [OH^-] = 1e-7 at this temperature. Furthermore, we define

pH = -log[H3O+]
pOH = -log[OH-]
pH = pOH = 7 at 298 K; (in neutral solutions)

It is important to realize that K_w depends on temperature as shown in the Table here.

Leveling Effect of Water and Acid-Base Characters

The strength of strong acids and bases is dominated by the autoionization of water. In aqueous solutions, the strongest acid and base are the hydronium ion, H_3O^+, and the hydroxide ion OH^- respectively. Acids HCl, HBr, HI, HNO_3, $HClO_3$, $HClO_4$, and H_2SO_4 completely ionize in water, making them as strong as H_3O^+ due to the **leveling effect** of water. Furthermore, strong acids, strong bases, and salts completely ionize in their aqueous solutions.

For example, HCl is a stronger acid than H_2O, and the reaction takes place as HCl dissolves in water.

HCl + H2O = Cl- + H3O+

A similar equation can be written for another strong acid.

On the other hand, a stong base also react with water to give the stong base species, OH^-.

H2O + B- = OH- + HB

For example, O^{2-}, CH_3O^-, and NH_3 are strong bases. The leveling effect also apply to bases.

Amphiprotic Species

Equilibria of acids and bases, are interesting chemistry. When an acid and a base differ by a proton, they are called a conjugate acid-base pair. A water molecule is a weak acid and base, due to its ability to accept or donate a proton. Such properties make water an **amphiprotic species**. In fact, H_3O^+, H_2O and OH^- are amphiprotic, as are some other conjugate acid-base pairs of weak acids and bases.

If several acids and bases are dissolved in water, all equilibria must be considered. To estimate the pH of these solutions requires the exact treatment of several equilibrium constants. For example, many species dissolve in rain water, and many equilibria must be considered. Detail consideration and examples are given in Acid-Base Reactions

Carbon dioxide in the air dissolve in rain water, lakes and rivers. A solution of CO_2 involves the following reaction:

Reaction	*K* formula	*K* value
$H_2O(l) + CO_2(g) = H_2CO_3(l)$	$1/P_{CO_2}$	?
$H_2CO_3 = HCO_3^- + H^+$	$[HCO_3^-]\,[H^+] / [H_2CO_3]$	5e-7
$HCO_3^- = CO_3^{-2} + H^+$	$[CO_3^{-2}]\,[H^+] / [HCO_3^-]$	5e-11
$HOH(l) + HOH(l) = H_3O^+ + OH^-$	$[H_3O^+]\,[OH^-]$	1e-14

These complicated equilibria make natural water a buffer.

Example 1

Assume that the partial pressure of carbon dioxide causes a total concentration of carbonic species to be 8e-4 M. Estimate the pH of this solution.

Solution

From the given data, we have the following five equations and five unknowns:

Equilibrium	Equations	No.
H_2CO_3 « $HCO_3^- + H^+$	$\frac{[HCO_3^-]\,[H^+]}{[H_2CO_3]} = 5e\text{-}7$	(1)
HCO_3^- « $CO_3^{2-} + H^+$	$\frac{[CO_3^{2-}]\,[H^+]}{HCO_3^-} = 5e\text{-}11$	(2)
2 H_2O « $H_3O^+ + OH^-$	$[H_3O^+]\,[OH^-] = 1e\text{-}14$	(3)
Charge balance	$[H^+] = [HCO_3^-] + [OH^-] + 2\,[CO_3^{2-}]$	(4)
All species containing C	$[H_2CO_3] + [HCO_3^-] + [CO_3^{2-}] = 8.0e\text{-}4$ M	(5)
Unknowns		
$[H^+]$, $[OH^-]$, $[H_2CO_3]$, $[HCO_3^-]$, $[CO_3^{2-}]$		

Solving these equations for the 5 unknowns can be done using Maple, Mathcad, spread sheet, or approximation. In any case, we are interested in the pH, and we can make the following approximations or assumptions

Assume H^+ mostly comes from (1)	$[H^+] = [HCO_3^-]$	
H_2CO_3 is a weak acid most unionize	$[H_2CO_3] = 8.0e\text{-}4$ M	(6)
Let $x = [HCO_3^-] = [H^+]$	$[HCO_3^-]\,[H^+] / [H_2CO_3] = x^2 / [H_2CO_3] = 5.0e\text{-}7$	(1)

Combining (1) and (6) gives $[H^+]^2 = x^2 = 8.0e\text{-}4 * 5.0e\text{-}7 = 4.0e\text{-}10$. Therefore,

$$[H+] = 2.0e\text{-}5$$
$$pH = -\log(2.0e5) = 4.7$$

DISCUSSION

Generally speaking, rain water has a pH about 5, rather acidic. It dissolves limestone and marble readily. Due to the dissolved carbon dioxide, rain water is a buffer solution.

Increased carbon dioxide level forces an increase in dissolved carbon dioxide. Would this causes pH of rain water to decrease or increase? Justify your answer by giving the reasons.

Since $[H^+]$ = 2.0e-5, $[OH^-]$ = 5e-9, the amount of H^+ from ionization of water is also 5.0e-9, small with respect to 2.0e-5 from ionization of H_2CO_3. Similarly, the ionization from

HCO3- « CO32- + H+

is also small. Most of the C-containing species is H_2CO_3

H_2CO_3 is a weak acid, its ionization is small indeed.

Now, you may proceed to evaluate other concentrations: $[OH^-]$, $[HCO_3^-]$, and $[CO_3^{2-}]$

Reactivity of Water Towards Metals

Alkali metals react with water readily. Contact of cesium metal with water causes immediate explosion, and the reactions become slower for potassium, sodium and lithium. Reaction with barium, strontium, calcium are less well known, but they do react readily. Warm water may be needed to react with calcium metal, however. Many metals displace H^+ ions in acidic solutions. This is often seen as a property of acids.

Electrolysis of Water

The *enthalpy of formation* for liquid water, $H_2O(l)$, is -285.830 and that of water vapour is -241.826 kJ/mol. The difference is the heat of vaporization at 298 K. Liquid water and vapor entropies (*S*) are 69.95 and 188.835 kJ K^{-1} mol^{-1} respectively, (*see* Thermodynamic Data. These are entropies, not standard entropies of formation. The entropy of formation for water is obtained by,

DSof water = Sowater - SoH2 - 0.5 SoO2
= 69.95 - 130.68 - 0.5*205.14 (data from Thermodynamic Data)
= - 163.3 J K-1 mol-1
DGowater = DH - T DS (note H in kJ/mol and S in J/mol)
DGowater = -285.83 - 298.15 * 163.3/1000 = -237.13 kJ

The equilibrium constant and Gibb's energy are related,

DGo = - R T ln K
K = exp(- DGo / R T)
= 3.5e41 atm-3/2

This is a very large value for the formation of water,

H2 + 0.5 O2 = 0.5 H2O(l).

In other words, the reaction is complete, and the possibility of water dissociated into hydrogen and oxygen is very small. A negative value for DG^o indicates an exothermic reaction.

The Gibb's energy is the energy released other than pressure-volume work. This redox reaction to form water can be engineered to proceed in a Daniel cell. In this case, the energy is converted into electric energy according to this equation.

DGowater = - n F E = -237.13 kJ

where *n* is the number of electrons (= 2) in the redox equation, *F* is the Faraday constant (= 96485 C), and *E* is the potential of the Daniel cell. Thus,

$$E = -\frac{-237130\ \text{J}}{2*96485\ \text{C}}$$

= 1.23 V

Ideally, a reverse voltage of 1.23 V is required for the electrolysis of water. But in reality, a little over voltage is required to carry out the electrolysis to decompose water. Furthermore, pure water does not conduct electricity, and acid, base or salt is often added for the electrolysis of water. This link has a demonstration.

Example 2

In order to carry out the electrolysis of water, 1.50 V is applied. Assume the energy not converted to chemical energy is converted to heat. How much heat is generated for the electrolysis of 1 mole water?

Solution

Ideally, 1.23 V will be used for the electrolysis. Energy due to the over voltage of 1.50 - 1.23 = 0.27 V is converted to heat.

Heat = 0.27 V * 2 * 96485 C
= 52102 J
= 52 kJ

DISCUSSION
The excess energy can also be evaluated using

Heat = n F *1.50 - 237130

This problem also illustrates the principle of conservation of energy.

Questions

1. **For the reaction**

H2O(l) -> H2(g) + 0.5 OH2(g)

the equilibrium constant as shown earlier is 1/3.5e41 = 2.9e-43 $atm^{3/2}$. What is the partial pressure of H_2(g)?

Solutions

1. Hint: 2.6e-64

Skill -
Evaluate this value please!

Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

This page titled 2.2.2.8: Water Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Chung (Peter) Chieh.

2.2.2.9: Water Physics

Physical Properties of Water

Chemical and physical properties of water are often discussed together. These properties are fundamentals of many disciplines such as hydrology, environmental studies, chemical engineering, environmental engineering, civil engineering etc. They are of interest to chemists and physicists of course.

Here are some highlights of the physical properties of water. Pure liquid water has a high heat capacity of 4.182 J K^{-1} g^{-1}; it is a good heat conductor, but a poor electric conductor. It is a solvent for dissolving ionic and polar substances, but interact with non-polar substance weakly. Surface tension of water is rather high, and little quantities aggregate into drops rather than spread out as thin layers.

Hydrogen bonding contributes to many of the physical and chemical properties, such as the unusual but normally known freezing point of 273.15 K, and 373.15 K respectively. The critical temperature and pressure are 647.3 K and 220.5 bar (22050 kPa) and critical volume = 0.056 m^3 $kmol^{-1}$.

Because of the many applications of water, some more details on the properties of water are desirable. Thus, keeping track of water properties is of national interest. For example, the American Society of Mechanical Engineers (ASME) is such an organization. International co-operation on research and information exchange are more economical, and for thermal properties, The International Association for the Properties of Water and Steam (IAPWS) is set up. Canada is a member of this organization.

From the application point of view, the variations of the following properties as functions of temperature and pressure are required.

- Compressibility of steam and water as a function of pressure at various temperature
- Density of water as a function of temperature
- Viscosity of water as a function of temperature
- Enthalpy of water for various thermodynamic evaluations
- Molar volumes and expansion coefficients of water and vapor as functions of pressure and temperature
- Speed of sound in water and vapor and speed of sound in air-vapor mixture as functions of temperature and pressure
- Entropy of water as a function of temperature and pressure
- Thermal conductivity of water and steam
- Viscosity of water at any temperature for pipe and pump design
- Dielectric constant as functions of temperature and pressure
- Surface tension as functions of temperature and pressure
- Gibbs energy at various temperature and pressure
- Properties such as dielectric constants and ion products of supercritical water (fluid)
- Applications of supercritical water: Plastic waste recycle, recovery of toluenediamine, hydrolysis of PET (polyethylene terephthalate) etc.

We will discuss some of these to illustrate the point, but not all of them.

Density of Water

Density is the mass per unit volume. The density of water is usually taken as 1.0 g/mL or 1.000e3 kg m^{-3} at 277 K. This suggests that the density varies with temperature and water density is the highest at 277 K, and the density between 273 and 281 K from the *CRC Handbook of Chemistry and Physics* are given in the Table here. These data are calculated from experimental data for pure water based on the standard at 276.98 K. The same source gives the density of ordinary water as 1.000000 g/mL at 277 K.

The volume occupied by one mole of substance is called the **molar volume**. The molar volume of liquid water is 18.016/*density*. At 277 K, the molar volume is 18.016 mL. For liquid water, the molar volumes of liquid water increase to 18.03 mL at both 269 K and 285 K.

The density of ice is 0.917 at 273 K, and the molar volume is 19.65 mL, 9% more than the molar of liquid. Thus, 9% of an ice cube containing no air bubble float above the surface, and 91% of it is below the waterline. The density makes behavior of icebergs interesting. Ice bergs are major tourist attractions in Newfoundland and Labrador, Canada.

Density of Water		
T /K	Density g/mL	D_2O
273	0.999841	1.10469
274	0.999900	
275	0.999941	
276	0.999965	
277	0.999973	1.1057
278	0.999965	1.10562
279	0.999941	
280	0.999902	
281	0.999849	
282	0.999781	
281	0.999700	

Electric Dipole Moment and Dielectric Constant

Charged ions interact with each other due to electrostatic attraction or repulsion. The force F between two charge particles with charges q_1, and q_2 separated by a distance r is

$$F = \frac{q1\ q2}{4\ p\ eo\ r2}$$

Cl d-
|
Na d+

Uncharged molecules still interact with each other, not due to electrostatic interaction, but due to electric dipole interaction. The electric dipole moment is a vector due to uneven distribution of unlike charges. In diatomic systems, the magnitude of the electric dipole moment can be estimated as the difference between the Pauling electronegativities of the two atoms. For convenience, let us assume that centres of positive and negatives of Na-Cl are separated by a distance l, then the electric dipole moment, μ is

$$m = q\ l$$

Traditionally, the dipole moments of molecules have been tabulated in electrostatic units, in which case the charge of an electron is 4.80e-10 esu (= (1.6e-19 C) (3e9 esu/C)). In NaCl crystals, the distance between Na and Cl ions is 240 pm. If the NaCl molecule (in a gas) has the same distance between the ideallized ions, then the dipole moment is calculated below:

m = q l
= (1.60e-19) (240e-12 m)
= 3.84e-29 C m.

or in cgs-esu units
m = 4.8e-10 esu * 2.40e-8 cm
= 11.5e-18 esu cm

In the cgs-esu unit, 1e-18 esu cm is define a Debye (symbol D). Thus, we have

1 D = 1e-18 esu cm
= 3.355e-30 C m (from the calculation above)
m = 11.5 D for ideallized NaCl gas molecule
but
mobserved = 9 D in NaCl gas =3e-29 C m.

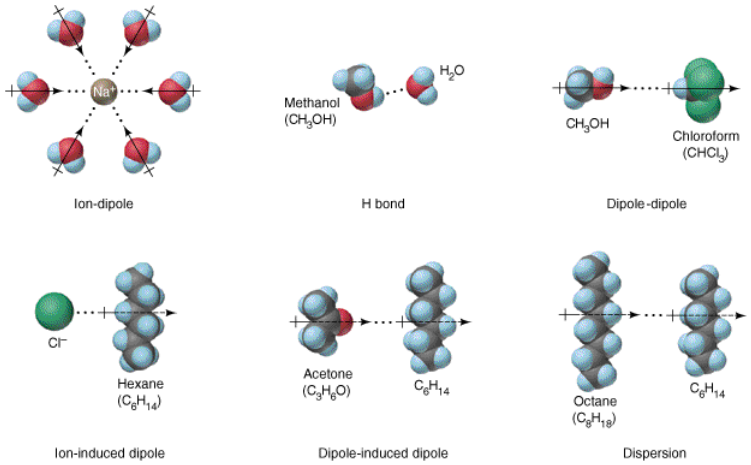
Electric Dipole Moment of Some Gas Molecules	
Molecule	m /D
NaCl	9.0
KCl	10.3
CO	0.1
HF	1.8
HCl	1.1
HBr	0.8
H_2O	**1.8**
SO_2	1.6
N_2O	0.2
NH_3	1.5

However, experimental dipole moment = 9 D for NaCl gas. Thus, the model for the calculation has to be modified to account for the partial delocalization of the charges or by including some covalent character in the Na-Cl bond. In any case, the model shows a physical method (model) for the evaluation of dipole moment. Dipole moments of some gas molecules are given in the table here.

The dipole moment is a vector from the negative to the positive charge along the bond. For triatomic moecules such as those of water, the total electric dipole moment is the sum of al dipoles for each bond. The experimental dipole moment for water is 1.8 D, which is the same as that of H-F. Water is a very polar compound. Ammonia with three N-H bonds has a dipole moment of 1.5 D.

Homonuclear diatomic molecules have zero dipole moment, of course. So are linear CO_2 and CS_2 molecules.

The high dipole moment makes water a very special substance. Water has a very high dielectric constant, 80. Due to dipole-charge interaction, water is the universal solvent for ionic substances, especially mono-valent ions. The diagram below shows some typical ion-dipole and dipole-dipole interaction in solutions. The dissolution in water is called hydration.



The dipole moment has something to do with its interaction with microwave. Application of Ground Penetrating Radar in Glaciology is a web site that gives the theory of radar and discusses the interaction of radar with water. It further illustrates the application in glaciology.

In contrast, supercritical water has a low dielectric constant, making it a good solvent for non-polar substances.

Example 1

Let us assume the H-F molecule as composed of two ions, a positive and a negative ion. What is the distance that separating these two ions in order to give a dipole moment of 1.8 D?

Solution

Recall the first formula from the discussion of dipole moment above, and proceed with the calculation as shown:

$$m = q\, l$$

$$q\, l = 4.8\text{e-}10 \text{ esu} * l$$

$$= 1.8\text{e-}18 \text{ esu cm.}$$

$$l = \frac{1.8\text{e-}18 \text{ esu cm}}{4.8\text{e-}10 \text{ esu}} = 0.375\text{e-}8 \text{ cm (37.5 pm)}$$

DISCUSSION

What is the normal H-F bond distance? (Ans. 92 pm).

Dividing the apparent bondlength of 38 pm by the observed bondlength 92 pm gives the ionic character 38/92 = 41 % ionic character.

Questions

1. **In a fire or nuclear power generator, steam at 600°C under pressure is often used to transfer the heat from the boiler to the turbine. What phase is this steam?**

Solutions

1. Hint: supercritical fluid of water.

Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

2.2.2.9: Water Physics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2.2.10: Water Treatment

Water treatment is` a process of making water suitable for its application or returning its natural state. Thus, water treatment required before and after its application. The required treatment depends on the application. For example, treatment of greywater (from bath, dish and wash water) differs from the black water (from flush toilets). Composting toilet is not allowed in urban dwelling. Yet, composting toilets are used in a 30,000-square-foot office complex at the Institute of Asian Research, University of British Columbia.

Water treatment involves science, engineering, business, and art. The treatment may include mechanical, physical, biological, and chemical methods. As with any technology, science is the foundation, and engineering makes sure that the technology works as designed. The appearance and application of water is an art.

In terms of business, RGF Environmental, Water Energy Technologies, Aquasana Store, Vitech, Recalyx Industrial SDN BHD and PACE Chemicals ltd are some of many companies that offer various processes for water treatment. Millipore, a Fisher Scientific partner, offers many lines of products to produce *ultrapure water*, using a combination of active charcoal membranes, and reverse osmosis filter. Internet sites of these companies offer useful information regarding water.

An environmental scientist or consultant matches the service provider, modify if necessary, with the requirement.

- Natural Water includes some discussion on hard and soft water. Softening hard water for boiler, cooler, and domestic application is discussed therein. These treatments prepare water so that it is suitable for the applications.
- Water Biology deals with water and biology. Drinking water is part of making water suitable for living. Thus, this link gives some considerations to drinking water problems.
- There are many different industry types, and waters from various sources are usually treated before and after their applications. Pre-application treatment and wastewater treatment offer a special opportunity or challenge. Only a general consideration will be given to some industrial processes.
- General municipal and domestic wastewater treatment converts used water (waste) into environmentally acceptable water or even drinking water. Every urban centre requires such a facility.

General Wastewater Treatment

Water is a renewable resource. All water treatments involve the removal of solids, bacteria, algae, plants, inorganic compounds, and organic compounds. Removal of solids is usually done by filtration and sediment. Bacteria digestion is an important process to remove harmful *pollutants*. Converting used water into environmentally acceptable water or even drinking water is **wastewater treatment**

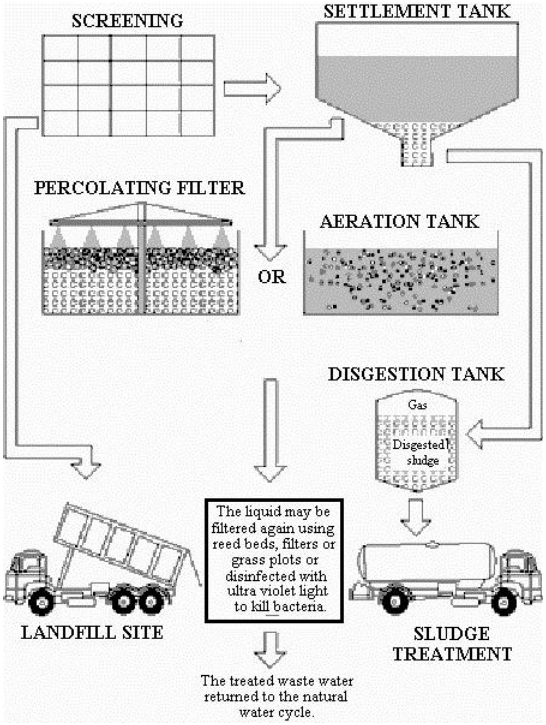
Water in the Great Lakes Region is an organization dealing with the water resources. Ontario Clean Water Agency (OCWA) is a provincial Crown corporation in business to provide environmentally responsible and cost-efficient water and wastewater services. It currently operates more than 400 facilities for 200 municipalities. This web site provides information on water and water treatment.

In April 1993, 403,000 people in Milwaukee were ill as a result of *cryptosporidium* contaimination of water due to spring run off. This outbreak caused the more stringent regulations to be implemented in the public dringking water system. The measures were aimed at removing cryptosporidium.

In May 2000, due to torrential downpour surface water got into shallow wells in a small town Walkerton, Ontario, Canada. On May 17, some residents complained of fever, bloody diarrhea and vomiting. This was know as the Walkerton *E. Coli* Outbreak. Nearly half of the population of the town fell ill, and several people died due to the E. Coli O157:H7 infection. A public inquiry recommended many measures to prevent similar outbreaks. These measures were aimed at eliminating E. Coli.

Sewage Treatment

As a general discussion, let us look at a typical process in sewage treatment. A flow diagram for a general sewage treatment plant from Water Education, Department of Computer Science, University of Exeter, U.K., is shown below:



Sewage is SCREENED to remove large solid chunks, which are disposed in LANDFILL SITE. It flows over to the SETTLEMENT TANK to let the fine particles to settle. The settlement is called the activated SLUDGE. The supernatant is then PERCOLATING FILTERED and/or AERATED. The water can be filtered again, and then disinfected (chlorinated in most cases). When there is no other complication, the water is returned to nature back to the ecological cycle.

The SLUDGE removed from the settlement is composed of living biological material. A portion of it may be returned to the AERATION TANK, but the raw SLUDGE is digested by both microorganism. Anaerobic (without oxygen) and aerobic (with air) bacteria digestions are used. At the digestion stage, carbon dioxide, ammonia, and methane gases are evolved. Volume of the digested sludge is reduced, and it is acceptable as a fertilizer supplement in farming.

Wastewater Treatment

Although the sewage water may be discharged back to the ecological system after AERATED DIGESTION and PERCOLATING FILTRATION, but in some cases, further treatment is required. Some general consideration of water treatment is given below.

A rather recent book, *Chemistry of Water Treatment* by S.D. Faust and O.M. Aly, 2nd Ed. (1998) [TD433 F38 1998], addresses the problem of quality natural and treated water.

The first three chapters discuss the criteria and standards for drinking water quality, organic compounds in waters, taste and order of water. Understandably, the standards change over the years. So are the standards of treated waters. Guidelines are available from government agencies such as Environment Canada which is equivalent to U.S. Public Health Service and the Environment Protection Agency (EPA). We have talked about drinking water in Water Biology.

Next seven chapters deal with the removal of the following:

- organics and inorganics by activated carbon
- particulate matter by coagulation
- particulate matter by filtration and sedimentation
- hardness and other scale-forming substances
- inorganic contaminants
- corrosive substances
- pathogenic (disease producing) bacteria, viruses, and protozoans (microorganisms).

There is a chapter dealing with **aeration**

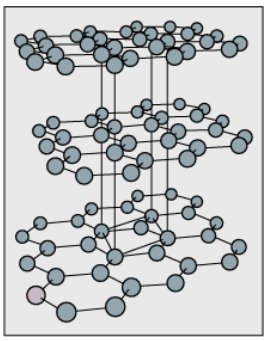
These items cover the chemistry, biology, and physics involved in the treatment of water. Some of these topics have been discussed in chemistry of water, physical properties of water, biology of water, and natural water. Introductions are going to be given to some selected topics below.

Treatment by Activated Carbon

Treatment by activated carbon is mostly due to adsorption or absorption. When a chemical species is adhered to the surface of a solid, it is an **adsorption**. When partial chemical bonds are formed between adsorbed species or when the absorbate got into the channels of the solids, we call it **absorption**. However, these two terms are often used to mean the same, because to distinguish one from type from the other is very difficult.

Application of activated charcoal for the removal of undesirable order and taste in drinking water has been recognized at the dawn of civilization. Using bone char and charred vegetation, gravel, and sand for the filtration of water for domestic application has been practised for thousands of years. Active research and production of activated charcoal was accelerated during the two world wars. The use of poison gas prompted the development of masks. They are still in use today.

Charcoal absorbs many substances, ranging from colored organic particulates to inorganic metal ions. Charcoal has been used to remove the colour of raw sugar from various sources.



Charcoal consists of microcrystallites of graphite. The particles are so small in charcoal that they were considered amorphous. The crystal structure of graphite consists of layers of hexagonal networks, stacked on top of each other. Today, making activated carbon is a new and widely varied industry. Other molecules attach themselves to the porous surface and dangling carbons in these microcrystallites.

Carbon containing substances are charred at less than 900 K to produce carbon in the manufacture of **activated carbon**. However, the carbon is activated at 1200 K using oxidizing agent to selectively oxidize portions of the char to produce pores in the material. Because of the special process to produce used, these materials with high surface to mass ratio, they are called **activated carbon** rather than activated charcoal. Factors affecting the absorption are particle size, surface area, pore structure, acidity (pH), temperature, and the nature of the material to be absorbed. Usually, adsorption (absorption) equilibria and rate of adsorption must be considered for effective applications.

Coagulation, Flucculation and Sedimentation

Natural and wastewater containing small particulates. They are suspended in water forming a **colloid**. These particles carry the same charges, and repulsion prevents them from combining into larger particulates to settle. Thus, some chemical and physical techniques are applied to help them settle. The phenomenon is known as **coagulation**. A well known method is the addition of electrolyte. Charged particulates combine with ions neutralizing the charges. The neutral particulates combine to form larger particles, and finally settle down.

Another method is to use high-molecular-weight material to attract or trap the particulates and settle down together. Such a process is called **flocculation**. Starch and multiply charged ions are often used.

Historically, dirty water is cleaned by treating with alum, $Al_2(SO_4)_3 \cdot 12\ H_2O$, and lime, $Ca(OH)_2$. These electrolytes cause the pH of the water to change due to the following reactions:

$$Al2(SO4)3.12\ H2O, \rightarrow Al(aq)3+ + 3\ SO4(aq)2- + 12\ H2O$$
$$SO4(aq)2- + H2O \rightarrow HSO4(aq)- + OH- \text{ (causing pH change)}$$
$$Ca(OH)2 \rightarrow Ca(aq)2+ + 2\ OH- \text{ (causing pH change)}$$

The slightly basic water causes $Al(OH)_3$, $Fe(OH)_3$ and $Fe(OH)_2$ to precipitate, bringing the small particulates with them and the water becomes clear. Some records have been found that Egyptians and Romans used these techniques as early as 2000 BC.

Suspension of iron oxide particulates and humic organic matter in water gives water the yellow muddy appearance. Both iron oxide particulates and organic matter can be removed from coagulation and flocculation. The description given here is oversimplified, and many more techniques have been applied in the treatment of water. Coagulation is a major application of lime in the treatment of wastewater.

Other salts such as iron sulfates $Fe_2(SO_4)_3$ and $FeSO_4$, chromium sulfate $Cr_2(SO_4)_3$, and some special polymers are also useful. Other ions such as sodium, chloride, calcium, magnesium, and potassium also affect the coagulation process. So do temperature, pH, and concentration.

Disposal of coagulation sludge is a concern, however.

Sedimentation let the water sit around to let the floculated or coagulated particles to settle out. It works best with relatively dense particles (e.g. silt and minerals), while flotation works better for lighter particles (e.g. algae, color). A settling tank should be big enough so that it takes a long time (ideally 4 hours +) to get through. Inlets and outlets are designed so the water moves slowly in the tank. Long and narrow channels are installed to let the water to snake its way through the tank. The settled particles, sludge, must occasionally be removed from the tanks. The water is next ready to be filtered. Sedimentation is used in pre-treatment and wastewater treatment.

Filtration

Filtration is the process of removing solids from a fluid by passing it through a porous medium. Coarse, medium, and fine porous media have been used depending on the requirement. The filter media are artificial membranes, nets, sand filter, and high technological filter systems. The choice of filters depends on the required filtering speed and the *cleanness* requirement. The flow required for filtration can be achieved using gravity or pressure. In pressure filtration, one side of the filter medium is at higher pressure than that of the other so that the filter plane has a pressure drop. Some portion of this filter type must be enclosed in a container.

The process of removing the clogged portion of the filter bed by reversing the flow through the bed and washing out the solid is called **back washing**. During this process, the solid must be removed out of the system, but otherwise the filters must be either replaced or taken out of service to be cleaned.

Aqua-Rain manufacture water filters as shown here. This unit consists of four filters. Regarding the filtering system, its technical info gave the following statement.



> *At the heart of the AquaRain™ Water Filtration System are Marathonr State-of-the-art ceramic elements utilizing a long-proven filtration process that is over 100 years old which will safely remove dangerous waterborne pathogens such as cysts (Cryptospordium, Giardia lamblia) and bacteria (E. coli, Samonelli typhus, etc...). These innovative Marathonr ceramic elements are also filled with a high grade silvered granulated activated carbon (GAC). The GAC reduces pesticides, chemicals, chlorine, tastes & odors, while leaving the naturally occurring minerals found in the water unaffected.*

The units are designed for emergency and perhaps undeveloped countries.

AquaSelect of Mississauga has a pitcher water filter system, and *its cartridge contains hundreds of high efficiency activated carbon and ion exchange beads*, its web site claims. Brita filters is very popular.

Aeration

Bringing air into intimate contact with water for the purpose of exchanging certain components between the two phases is called **aeration**. Oxygenation is one of the purposes of aeration. Others are removal of volatile organic substances, hydrogen sulfide, ammonia, and volatile organic compounds.

A gas or substance dissolved in water may further react with water. Such a reaction is called **hydration**. Ionic substance dissolve due to hydration, for example:

$$\mathrm{HCl(g)} + x\,\mathrm{H_2O} \rightleftharpoons \mathrm{H(H_2O)}_x{}^{+} + \mathrm{Cl(aq)}^{-} \tag{2.2.2.10.1}$$

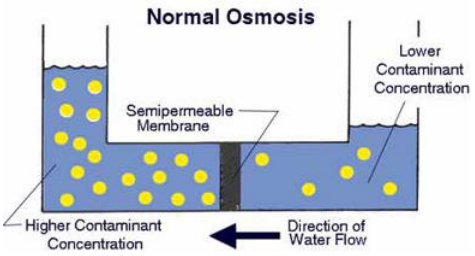
$$\mathrm{H_2S} \rightleftharpoons \mathrm{H^{+}\,(aq)} + \mathrm{HS^{-}\,(aq)} \tag{2.2.2.10.2}$$

These reactions are reversible, and aeration may also causes dehydration resulting in releasing the gas from water. Henry's law is applicable to this type of equilibrium for consideration. Methods of aeration are

- Diffused aeration - Air bubbles through water.

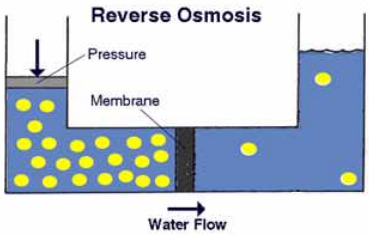
- Spray aeration - Water is sprayed through air.
- Multiple-tray aeration - Water flows through several trays to mix with air.
- Cascade aeration - Water flows downwards over many steps in the form of thin water falls.
- Air stripping - A combination of multiple tray and cascade technique plus random packed blocks causing water to mix thoroughly with air.

Reverse Osmosis Water Filter System



In the following discussion, a dilute solution and a concentrated solution are considered. The dilute solution can be a clean water whereas the concentrated solution contains undesirable solute (electrolyte or others).

When a compartment containing a dilute solution is connected to another compartment containing a concentrated solution by a semipermeable membrane, water molecules move from the dilute solution to concentrated solution. This phenomenon is called osmosis. Pig bladders are natural semipermeable membranes. As the water molecules migrate through the semipermeable membrane, water level in the solution will increase until the (osmotic) pressure prevents a net migration of water molecules in one direction. A pressure equivalent to the height difference is called the osmotic pressure. The illustration given on the right is from the PurePro, one of the many companies that manufacture reverse osmosis water filter devices. Millipore also use this technique.



By applying pressure in the higher concentration solution, water molecules migrate from a high concentration solution to a low concentration solution. This method is called **reverse osmosis water filter system**. The concept of reverse osmosis is illustrated in the diagram here from PurePro.

In this technique, the membrane must be able to tolerate the high pressure, and prevent solute molecules to pass through. Regarding membranes, PurePro made the following statement:

> *Semipermeable membranes have come a long way from the natural pig bladders used in the earlier osmosis experiments. Before the 1960's, these membranes were too inefficient, expensive, and unreliable for practical applications outside the laboratory. Modern advances in synthetic materials have generally solved these problems, allowing membranes to become highly efficient at rejecting contaminants, and making them tough enough to withstand the greater pressures necessary for efficient operation.*

This technology certainly works, and it has been used to convert salt (ocean or sea) water into fresh water. With this technique, the water with higher concentration is discharged. Thus, this technology is costly in regions where the water cost is high. Free Drinking Water also uses reverse osmosis filter system for domestic applications.

Industrial Wastewater Treatment

The Environment Canada's (Atlantic Region) Waste Management and Remediation Section contains links to detailed information on the programs and activities relating to Petroleum and Allied Petroleum Storage Tank Systems, Ocean Disposal, Contaminated Sites, and Hazardous Waste Disposal Advice for the provinces of Newfoundland and Labrador, Nova Scotia, New Brunswick and Prince Edward Island. It also contains links to sites within and outside of Environment Canada with related information.

Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

This page titled 2.2.2.10: Water Treatment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Chung (Peter) Chieh.

SECTION OVERVIEW

2.2.3: Atmospheric Chemistry

Atmospheric chemistry is a branch of atmospheric science in which the chemistry of the Earth's atmosphere and that of other planets is studied.

2.2.3.1: Atmosphere

2.2.3.2: Carbon Cycle

2.2.3.3: Carbon Oxides

2.2.3.4: Ozone

2.2.3.5: Photosynthesis

Contributors and Attributions

- Template:Contribchieh

2.2.3: Atmospheric Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2.3.1: Atmosphere

Discussion Questions

- How do atmospheric scientists view the atmosphere?
- What gases are pollutants in the atmosphere?

The atmospheric chemistry studies the chemical composition of the natural atmosphere, the way gases, liquids, and solids in the atmosphere interact with each other and with the earth's surface and associated biota, and how human activities may be changing the chemical and physical characteristics of the atmosphere. It is interesting to note that the 1995 Nobel Prize in Chemistry 1995 was awarded to the atmospheric scientists P. Crutzen, M. Molina and F. S. Rowland. For convenience of study, atmospheric scientists divide the atmosphere as if it consists of four layers. The division is mainly due to temperature variations as the altitude increases. The four layers according to the variation of temperature are.

- Ionosphere (Aurora) or Thermosphere
- Mesosphere
- Stratosphere
- Troposphere

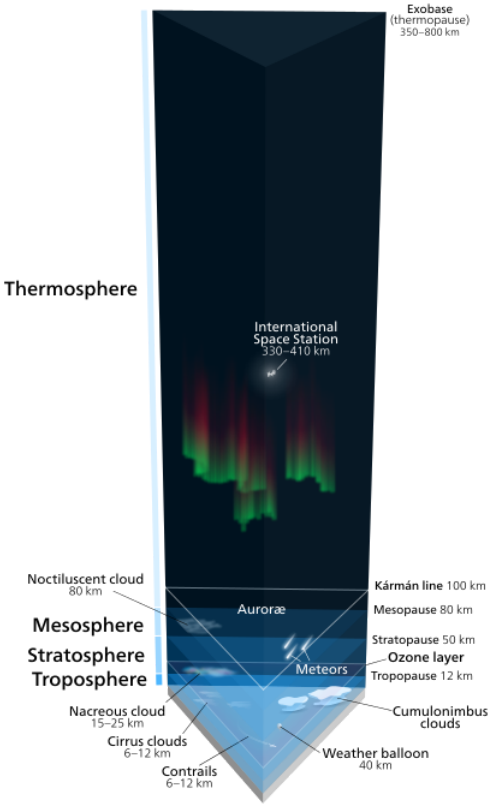


Figure 2.2.3.1.1: Earth's atmosphere Lower 4 layers of the atmosphere in 3 dimensions as seen diagonally from above the exobase. Layers drawn to scale, objects within the layers are not to scale. Aurorae shown here at the bottom of the thermosphere can actually form at any altitude in this atmospheric layer. from NASA.

Above 100 km is the thermosphere and ionosphere where the temperature increases from 200 K at 100 km to 500 K at 300 km. The temperature goes even higher as the altitude increases. activity as the altitude decrease. In the outer space, most particles consist of single atoms, H, He, and O etc. At lower altitude (200 - 100 km), diatomic molecules N_2, O_2, NO etc are present. The ionosphere is full of electrically charged ions. The UV rays ionizes these gases. The major reactions are

In the ionosphere:

$$O + hv \rightarrow O^+ + e^- \tag{2.2.3.1.1}$$

$$N + h\nu \rightarrow N^+ + e^- \tag{2.2.3.1.2}$$

In the neutral thermosphere:

$$N + O_2 \rightarrow NO + O \tag{2.2.3.1.3}$$

$$N + NO \rightarrow N_2 + O \tag{2.2.3.1.4}$$

$$O + O \rightarrow O_2 \tag{2.2.3.1.5}$$

Beyond the neutral thermosphere is the ionosphere and exosphere. These layers are of course interesting for space explorations and environmental concerns and space sciences. The atmosphere in the outer space is more like a plasma than a gas. Below the thermosphere is the mesosphere (100 - 50 km) in which the temperature decreases as the altitude increase. In this region, OH, H, NO, HO_2, O_2, and O_3 are common, and the most prominent chemical reactions are:

$$H_2O + h\nu \rightarrow OH + H \tag{2.2.3.1.6}$$

$$H_2O_2 + O \rightarrow OH + OH. \tag{2.2.3.1.7}$$

Below the mesosphere is the **stratosphere**, in which the temperature increases as the altitude increase from 10 km to 50 km. In this region, the following reactions are common:

$$NO_2 \rightarrow NO + O \tag{2.2.3.1.8}$$

$$N_2O \rightarrow N_2 + O \tag{2.2.3.1.9}$$

$$H_2 + O \rightarrow OH + H \tag{2.2.3.1.10}$$

$$CH_4 + O \rightarrow OH + CH_3 \tag{2.2.3.1.11}$$

Air flow is horizontal in the stratosphere. A thin ozone layer in the upper stratosphere has a high concentration of ozone. This layer is primarily responsible for absorbing the ultraviolet radiation from the sun. The ozone is generated by these reactions:

$$O_2 + h\nu \rightarrow O + O \tag{2.2.3.1.12}$$

$$O_2 + O \rightarrow O_3 \tag{2.2.3.1.13}$$

The troposphere is where all weather takes place; it is the region of rising and falling packets of air. The air pressure at the top of the troposphere is only 10% of that at sea level (0.1 atmospheres). There is a thin buffer zone between the troposphere and the next layer called the tropopause.



Figure 2.2.3.1.2: Space Shuttle Endeavour silhouetted against the atmosphere. The orange layer is the troposphere, the white layer is the stratosphere and the blue layer is the mesosphere.[1] (The shuttle is actually orbiting at an altitude of more than 320 km (200 mi), far above all three layers.). from NASA.

The major components in the region close to the surface of the Earth are N_2 (78%), O_2 (21%), Ar (1%) with variable amounts of H_2O, CO_2, CH_4, NO_2, NO_2, CO, N_2O, and O_3. The ozone concentration in this layer is low, about 8% of the total ozone in the atmosphere is in the troposphere.

What gases are pollutants in the atmosphere?

From the atmospheric science viewpoint, interactions of all gasses among themselves and their interaction with the environmental elements are of interest. However, for identification purposes, we need to identify the gases produced by man-made process (industry).

Some of the gases due to human activities are:

- Carbon dioxide result from the excess burning of carbon-containing fuel.
- Carbon monoxide produced by automobiles. This orderless and colorless gas is very toxic.
- Ozone produced in the exhaust of internal combustion engine, and the variation of ozone concentration in the stratosphere.
- Nitrogen oxides such as NO, NO_2, N_2O_4; due to the production of NO in the internal combustion engine.
- Methane gas produced due to treatments of large amount of waste.
- Sulfur oxides produced in mining operation and in the combustion of sulfur containing fuel. Sulfur oxide causes the so called *acid rain* problem.
- Chlorofluorocarbons (CFC) are gases used as refrigerant. When disposed into the atmosphere, they cause the ozone concentration to decrease.

Water vapor is also considered a greenhouse gas, but it is also generated by nature continuously due to radiation from the Sun. Of course, when water vapor condense into a liquid, much energy is released in the exothermal process. Condensation of water vapor causes storms and many of the weather phenomena.

Questions

1. According to what is the atmosphere divided into 4 layers?

   Skill -

   Describe the structure of the atmosphere.
2. Which layer contains the most ozone?

   Skill - Describe all the details of ozone?
3. How thick is the troposphere?

   Discussion - At the top of the world highest mountain, ~10 km in altitude, the atmosphere is only 0.1 of that at sea level. This is the top of the troposphere.
4. What type of gas is present in the thermosphere?

   Skill - Explain the chemistry taken place in the thermosphere?
5. How is the ionosphere different from other layers?

   Discussion - The aurora is related to the ions in the atmosphere.
6. What causes the gas molecules in the ionosphere to ionize and become charged particles?

   Skill - Describe the chemistry in the ionosphere.
7. If ozone is a beneficial gas in the atmosphere, why is ozone also a gaseous pollutant?

   Discussion - Decomposition of ozone releases O, OOH, OH radicals and they are harmful to many living organisms.
8. Why are chlorofluorocarbons a gases pollutant?

   Discussion - Ozone in the stratosphere absorbs harmful UV C and UV B, which are harmful to humans and plants.
9. Why is warming up of the Earth a bad thing?

   Skill - Give your opinion please on an issue.

Solutions

1. The division is made according to patterns of temperature variation.
2. The stratosphere, between 15 and 50 km.
3. The troposphere ranges from 8 to 15 km.
4. A very dilute concentration of monoatomic gas.
5. The ionosphere contains high concentration of charged particles.
6. Radiation or high energy photons from the sun caused ionization of atoms.
7. Ozone is very reactive, causing harm to living organisms.
8. Because they catalyze the decomposition of ozone in the stratosphere.
9. Because we do not know what is the consequence in the future.

Learning Guide

- How is the atmosphere divided into layers? What are the names of the layers?
- Where is the ozone layer located in the atmosphere? What is the molecular structure of ozone? How is the ozone formed?

Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

This page titled 2.2.3.1: Atmosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Chung (Peter) Chieh.

- **Atmosphere** by Chung (Peter) Chieh is licensed CC BY-NC-SA 4.0.

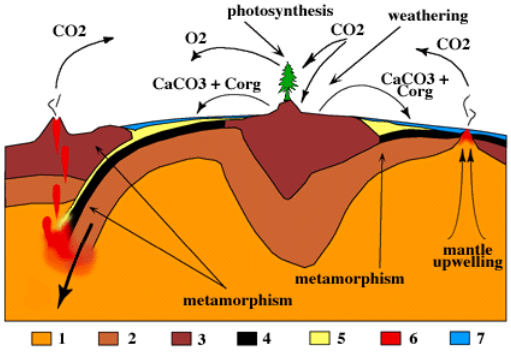
2.2.3.2: Carbon Cycle

Discussion Questions

- How is carbon cycled at a global scale?

The Global Carbon Cycle

A scientist and an engineer may be called upon to solve a particular problem involving coal (carbon), gasoline (hydrocarbon), combustion of carbon or carbon containing fuel, lime stone, sea shells, carbon monoxide, or carbon dioxide. When we formulate a solution, we should be aware of the impact not only of the problem, but also of the solution for such a problem. Otherwise, the solution may result in a problem that is more expensive to solve later. Thus, it is important to know how carbon evolve at a global scale. The carbon cycle? is part of the Earth cycle. The diagram from this link is shown here, because it illustrate the global cycle of carbon without including respiration and metabolism. It illustrates the geological processes.



How is carbon cycled at a global scale?

The carbon atoms undergo a complicated chemistry forming what is known as the global carbon cycle, as do oxygen, nitrogen, and other elements; but the carbon cycle is the most widely recognized. An animal produces carbon dioxide and consumes oxygen in its metabolism of food. Glucose is a typical food and a metabolic reaction can be represented by:

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O \tag{2.2.3.2.1}$$

A plant and green bacteria, on the other hand, produces oxygen and consumes carbon dioxide in its photo synthesis. Energy in the form of electromagnetic radiation (or photons) is supplied so that the low-energy-content carbon dioxide can be converted to high-energy-content glucose. An overall reaction for the complicated multi-step photosynthesis reaction can be represented by:

$$6CO_2 + 12H_2O - -h\nu \rightarrow C_6H_{12}O_6 + 6O_2 + 6H_2O \tag{2.2.3.2.2}$$

At a glance, animals and plants make food for each other. The plants convert solar energy into high-energy food for the animals. Water is a reactant and a product in the photosynthesis. Radioactive labeled studies showed that the oxygen in the water produced comes from those in carbon dioxide. You may be thinking about plants with leaves that give beautiful flowers. In fact, primitive plants in the ocean play a more important role in in the photosynthesis process, because of the large number of them.

The solubility of carbon dioxide depends on its partial pressure. As we know, carbon dioxide dissolves in water to form carbonic acid:

$$CO_2 + H_2O \rightarrow H_2CO_3 \tag{2.2.3.2.3}$$

$$H_2CO_3 \rightarrow H^+ + HCO_3^- K_{a1} = 4.2 \times 10^{-7} \tag{2.2.3.2.4}$$

$$HCO_3^- \rightarrow H^+ + CO_3^{2-} K_{a2} = 4.8 \times 10^{-11} \tag{2.2.3.2.5}$$

The dissolved carbon dioxide further reacts with metal ions in the water forming calcium and magnesium carbonates. The K_{sp} values for $CaCO_3$ and $MgCO_3$ are 5×10^{-9} and 3×10^{-3} respectively. Extensive limestone ($CaCO_3$) and dolomite (mixture of $CaCO_3$ and $MgCO_3$) have been formed this way.

$$CaCO_3 \rightarrow Ca^{2+} + CO_3^{2-} \quad K_{sp} = 5 \times 10^{-9} \tag{2.2.3.2.6}$$

$$MgCO_3 \rightarrow Mg^{2+} + CO_3^{2-} \quad K_{sp} = 3 \times 10^{-3} \tag{2.2.3.2.7}$$

Some believed that this is how lime stone produced. Lime stone is soluble in acidic solutions, which may be formed by dissolving large amount of carbon dioxide.

$$CaCO_{3(s)} + 2H^+_{(aq)} \rightarrow Ca^{2+}(aq) + H_2CO_{3(aq)} \tag{2.2.3.2.8}$$

or,

$$CO2(aq) + H2O(l) + CaCO3(s) \rightarrow Ca2 + (aq) + 2HCO3 - (aq) \tag{2.2.3.2.9}$$

When the concentration of carbon dioxide is reduced, the acidity decreases and the reverse reaction takes place forming a solid, $CaCO_3$(s). Thus, metabolism, photosynthesis, mineralization and geological process are the major chemical processes in the global carbon cycle.

Example 1

In general, it is known that rain water saturated with carbon dioxide has a pH of 5.6. Lower than 5.6 is called acid rain due to the presence of sulfur oxides and nitrogen oxides. Assume the water to be otherwise pure than the dissolved carbon dioxide, estimate the solubility of carbon dioxide in water?

Solution

Since pH = 5.6,

$$[H+] = 10 - 5.6M$$

$$= 2.5 \times 10^{-6} M$$

Thus, the contribution of hydrogen ions from self ionization of water (pH = 7) is negligible. We have $[H^+] = 2.5 \times 10^{-6}\ M = [HCO_3^-]$

The ionization of dissolved carbon dioxide is represented by these reactions,

$$CO_2 + H_2O \rightarrow H_2CO_3$$

$$H_2CO_3 \rightarrow H^+ + HCO_3^- \quad K_{a1} = 4.2 \times 10^{-7}$$

$$HCO_3^- \rightarrow H^+ + CO_3^{2-} \quad K_{a2} = 4.8 \times 10^{-11}$$

The major contribution to the production of hydrogen ion comes from the first ionization of H_2CO_3, and other contributions are almost negligible. If we assume the concentration of H_2CO_3 to be *x* M in its ionization,

$$H_2CO_3 \rightarrow H^+ + HCO_3^- \quad K_{a1} = 4.2 \times 10^{-7}$$

then, by definition of K_{a1} we have

$$\frac{(2.5 \times 10^{-6})}{x} = 4.2 \times 10^{-7}$$

Thus,

$$x = 1.5 \times 10^{-5}\ M$$

$$= 0.65 \text{ mg / L}$$

Thus, the solubility is about 0.65 ppm by weight.

DISCUSSION

Many assumptions have been made here, other wise the solution will be more complicated. Make sure you understand the

assumptions.

Example 2

From the solubility products of $CaCO_3$, estimate its molar solubility in natural water saturated with carbon dioxide at pH 5.6 and 298 K.

Solution

From the estimates given in Example 1, we still have to consider these equilibria:

$$H_2CO_3 \rightarrow H^+ + HCO_3^- \quad K_{a1} = 4.2 \times 10^{-7}$$

$$HCO_3^- \rightarrow H^+ + CO_3^{2-} \quad K_{a2} = 4.8 \times 10^{-11}$$

Since the second ionization constant K_{a2} << K_{a1}, it is safe to assume the following:

$$[H+] = [HCO3-] = 2.5 \times 10^{-6} M$$

in the above equilibria. The following formulation is derived by the definition of K_{a2}:

[H^+] [CO_3^{2-}] 2.5\times 10^{-6 [CO_3^{2-}]
---------- = ---------- = 4.8\times 10^{-11}
[HCO_3^-] 2.5\times 10^{-6

Thus, [CO_3^{2-}] = 4.8\times 10^{-11.

By the definition of the solubility product,

$$CaCO_3 \rightarrow Ca^{2+} + CO_3^{2-} \quad K_{sp} = 5 \times 10^{-9} \qquad (2.2.3.2.10)$$

we have

[Ca2+] = 5\times 10^{-9/4.8\times 10^{-11
= 100 M

DISCUSSION

This value is obviously too high and unreasonable. The result is certainly incorrect. Thus, we should re-examine the last assumption. As $CaCO_3$ dissolves, the concentration of carbonate ion also increase. If this concentration is high, then the contribution due to dissolved carbon dioxide is negligible. Effectively, we have [Ca^{2+}] = [CO_3^{2-}] = y, and

$$y^2 = 5 \times 10^{-9}$$

Thus,

y = 7\times 10^{-5 M = 0.007 g CaCO3 / L.

This results is obtained by ignoring the dissolved carbon dioxide. The true value is probably somewhere in between, because as calcium carbonates dissolves, the pH of the solution changes.

Questions

1. **Skill** - Explain the chemistry involved in the **global carbon cycle**.
2. **Skill** - Explain and identify the chemical composition of minerals and ores.
3. **Skill** - Explain the meaning of a scientific datum, K_{sp} in plain language.
4. **Skill** - Describe the trend of a chemical property when a variable changes.

Solutions

1. Metabolism, photosynthesis, and mineralization are the major chemical processes.
2. Calcium carbonate
3. Since Ksp for MgCO3 is larger, it is more soluble.

4. In general terms, the solubility of limestone decreases as pH increases.

Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

This page titled 2.2.3.2: Carbon Cycle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Chung (Peter) Chieh.

- **Carbon Cycle** by Chung (Peter) Chieh is licensed CC BY-NC-SA 4.0.

2.2.3.3: Carbon Oxides

Discussion Questions

- What are the molecular structures of carbon oxides?
- What atomic orbitals are involved in the molecular orbitals of carbon oxides?
- Why do CO molecules form strong bonds with metal atoms in carbonyls?
- What are some of the applications of carbon oxides?
- How has carbon dioxide level changed?
- What measure can be taken to reduce CO_2 emission?

Carbon Oxides

Carbon forms two important gases with oxygen: carbon monoxide, CO, and carbon dioxide, CO_2. Carbon oxides are important components of the atmosphere, and they are parts of the carbon cycle.

Carbon dioxide is naturally produced by respiration and metabolism, and consumed by plants in their photosynthesis. Since the industrial revolution, greater amount of carbon dioxide has been generated for over a hundred years due to increased industrial activities.

Today, information on carbon oxides is important. Issues related to carbon oxides have no boundaries. The Carbon Dioxide Information Analysis Center (CADIAC) provides global datasets on carbon dioxide and other atmosphere gases and climate. These datasets are available to international researchers, policymakers, managers, and educators to help evaluate complex environmental issues associated with potential climate change.

Carbon monoxide is also a national and global concern. The Consumer Product Safety Commission (CPSC) considers CO a senseless killer, and it provides information on CO poisoning and detection.

What are the molecular structures of carbon oxides?

The formation of carbon oxides is due to electronic configurations of carbon and oxygen. They have 4 and 6 valence electrons respectively. Using these valence electrons, we can give the Lewis dot structure for CO and three resonance structures for CO2 as follows:

```
           ..     ..     ..                ..
:C:::O:    :O::C::O: « :O:C:::O: « :O:::C:O:
                      ''                ''
```

These formulas suggest very strong bonding between carbon and oxygen in these gaseous molecules: triple bond in CºO, and double bonds in O=C=O. However, a formula containing a triple bond contribute to the resonance structure.

What atomic orbitals are involved in the molecular orbitals of carbon oxides?

The chemical bonding is more of an interpretation of the molecules in view of their properties. Using results from quantum mechanical approach, we may start by reviewing the electronic configurations of carbon and oxygen:

C: $1s^2\ 2s^2 2p^2$
O: $1s^2\ 2s^2 2p^4$

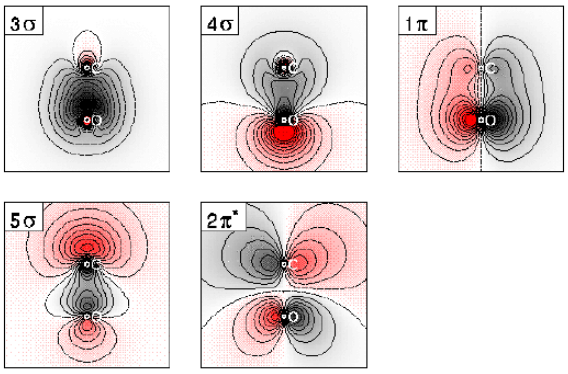
Thus, the carbon has 4 valence electrons and oxygen has 6 valence electrons. The s and p atomic orbitals are available for chemical bonding.

The valence bond approach suggests that p orbitals of carbon and oxygen are used in these molecules. In CO, only one such atomic orbital from each atom of C and O are employed to form a sigma, s, bond, and overlapping of two p orbitals leads to the formation of the two pi, p, bond. Thus, the bond order is 3 between C and O in CºO.

Note

A CO molecule has the same number of electrons as N_2, and these molecules are said to be **iso-electronic**. The N_2 molecule is also represented by NºN.

The molecular orbital (MO) approach for CO is describe in the lecture, and the MO energy level diagram has been given there. The plots of contours of equal electron densities has also been shown in earlier lectures, and the diagram for CO molecular orbitals is shown below:



Free CO molecular orbitals

The valence bond approach for CO_2 bonding is also very interesting. The two sp hybrid orbitals of central carbon overlaps with one p orbital from each of the oxygen atoms to from the two C-O s bonds in O-C-O. The two remaining p orbitals of carbon overlap with a p orbital each of the two oxygen atoms forming two p bonds, leading to the formation of O=C=O.

Here is a challenge: find a suitable diagram for either valence bond approach or for the MO approach for carbon dioxide in the web.

Why do CO molecules form strong bonds with metal atoms in carbonyls?

Molecules of CO and NN have two π bonds. Since the two atoms in CO are different, this made CO much more reactive than nitrogen. Indeed, CO forms many carbonyls with metal atoms or ions. For example, you have encountered some of the following carbonyls on the page of Heterogeneous catalysts

- Ni(CO)4
- Fe(CO)5
- Co2(u-CO)2(CO)6, (u-CO meaning CO bridged between two metal atoms)
- Mn2(CO)10
- Fe3(CO)12
- Co4(CO)12
- Rh4(CO)12
- CFe5(CO)15
- Rh6(CO)16
- Os6(CO)18

The study of metal carbonyls started with the discover of ferrocene $Fe(C_5H_5)_2$ and now hundreds if not thousands of metal carbonyls have been synthesized..

A sbond is formed between the carbon of CO and the metal atom. Such a bonding is very sensible if we consider the sp hybrid atomic orbitals of carbon being used in this case. Since there are two electrons in this orbitals of CO, the metal atom gains at least some fraction of electron due to the formation of this bond.

The empty antibonding pi orbital π^* of CO has the right symmetry and orientation to receive the back-donated electron from the filled d-orbitals of the metal atom. The back donation reinforce the sigma bond, and vice versa. This type of bonding has been

called the **synergic bonding mechanism** by Cotton and Wilkinson in their *Advanced Inorganic Chemistry*. A diagram showing this type of bonding scheme is shown on page 159 in *Inorganic Chemistry* by Swaddle.

What are some of the applications of carbon oxides?

Carbon oxides are useful commodities. A gas containing CO and hydrogen is called synthetic gas, because it can be converted to methanol using a catalyst. During the past few decades, many metal carbonyls have been prepared. These carbonyls are potential catalysts. When the metal carbonyl is a gas, the purified metal carbonyl gas can be used for the production of extra-pure metals.

Carbon dioxide is also a useful industrial gas. It is widely used in food and beverage industry. Here are some of its applications.

- making efferescent drinks
- manufacture of urea, $CO(NH_2)_2$, as fertilizer
- promote plant growth in green house
- making dry ice
- fire extinguisher
- provide an inert atmosphere for fruit and vegetable preservation
- as a supercritical fluid for solvent extraction

The critical temperature of carbon dioxide is only 304 K at a critical pressure of 7.39 MPa (almost 7 atm). These conditions can easily be met to generate supercritical carbon dioxide, which is a powerful and descriminating solvent. The supercritical fluid penetrates porous solids, evaporates without leaving a trace. Thus, this fluid is widely used as an extracting solvent. This fluid is also very useful in the field of analytical chemsitry.

On the high technology side, carbon dioxide lasers can provide a continuous laser beam from several milliwatts to several killowatts with a typical efficiency of 30%, one of the most efficient laser generation devices. This link also illustrates the basic theory of laser. Among many applications of laser, carbon dioxide laser has been used for skin resurfacing as an art of cosmetic surgery.

On the other hand, the heavier carbon dioxide usually stay in lower grounds, and when its concentration is very high, it can be a thread to living beings who will die of asphyxiation.

How has carbon dioxide level changed?

When Henry Ford put people to work on the assembly line, he did not worry about the consequence of automobile exhaust. He probably did not foresee the change of the society. Now everyone wants a piece of the carbondioxide generating machine. You can imagine that when all countries use as much energy as Canadians do, the carbon dioxide level of the atmosphere will be much higher. We need good measurements about carbon dioxide level in order to know how it is changing.

The National Oceanic and Atmospheric Administration (NOAA) of the U.S. is keeping track of it, and the measurements at Barrow, Alaska. The annual increase has been reported to be 1.49 ppm by volume per year. On the other hand, the atmospheric CO_2 concentration was about 280 ppm by volume in the 1700s before the industrial revolution, and it was 360 ppm in 1994. If you want more details about Carbon Dioxide Emissions in the U.S., this link is full of data.

Engineers, scientists, politician and the general public believe that increase level of CO_2 will cause the world to warm up, because some scientists have demonstrated their findings and the experts agreed. Lengthy discussion is required to present scientific evidences for the so called *green house effect* of CO_2, and hopefully some day you will be able to judge the argument yourself. I have not found simple and convincing evidence to present at this time. However, experts have suggested a correlation with the increase level of CO_2 and the average temperature of the globe.

What measure can be taken to reduce CO_2 emission?

The green house effect of CO_2 has attracted attention not only of expert and politicians, the public pressure (mostly the news media) have made the reduction of CO_2 emission an international priority. The United Nations Panel on Climate Change made several recommendations regarding CO_2.

- use natural gas as fuel rather than coal or oil
- use solar or nuclear energy instead of fossil fuels for electricity generation

- reduce rate of deforestation
- limit use of automobiles

Well, what can be done at a personal level, as a community, and as a country requires the determination of individuals. This is a challenge for us all especially engineers, because they are at the forefront of many industries. We face many front for a solution to the problem of carbon dioxide emission.

Example 1

The standard enthalpy of formation of NO is 90.25 kJ / mol and the standard entropies of N_2, O_2, NO are 191.61, 205.138 and 210.761 J (K mol)$^{-1}$ respectively. Calculate the Gibb's energy for the reaction

$$\mathrm{N_2 + O_2 \rightarrow 2\,NO}$$

at standard conditions.

Solution

The entropy of formation for the above reaction is

$$\begin{aligned}\Delta S^o &= 2 \times 210.761 - (205.138 + 191.61) \\ &= 24.77 \frac{J}{Kmol}\end{aligned}$$

The Gibb's energy of formation is then:

$$\begin{aligned}\Delta G^o &= \Delta H^o - T\Delta S^o \\ &= 180.5 - 298 * 0.02477 \\ &= 180.5 - 7.38\,kJ \\ &= 173.12\,kJ\end{aligned}$$

This value shows that the reaction is endothermic.

DISCUSSION

What is the equilibrium constant for the reaction as written and what is the implication of the result in the discussion of NO in air?

Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

This page titled 2.2.3.3: Carbon Oxides is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Chung (Peter) Chieh.

- **Carbon Oxides** by Chung (Peter) Chieh is licensed CC BY-NC-SA 4.0.

2.2.3.4: Ozone

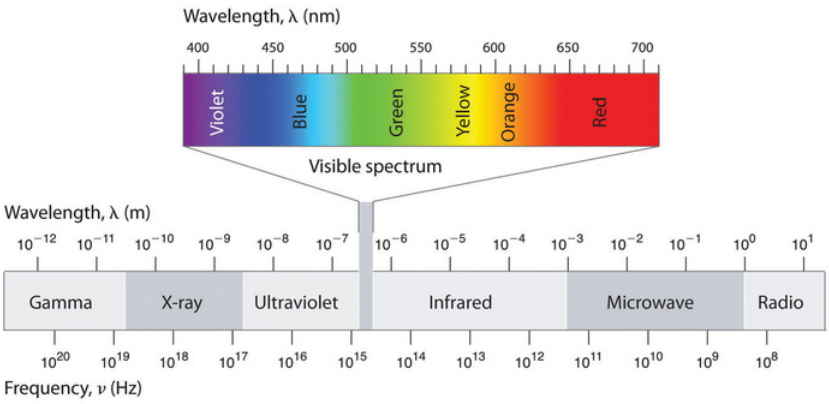
Discussion Questions

- What is UV?
- How is ozone produced in the atmosphere?
- How much ozone is in the atmosphere, and where is the ozone layer?
- What is the interaction of ozone and UV?
- What is ozone depletion?
- What is ozone hole and how does it vary over time?
- What are CFCs?
- How do CFCs help depleting ozone?
- How is ozone depletion in the polar region different from other regions?
- What has been done and what can be done to reduce ozone depletion?

Most of the ozone in the atmosphere is in the stratosphere of the atmosphere, with about 8% in the lower troposphere. As mentioned there, the ozone is formed due to photo reaction. The ozone level is measured in Dobson Unit (DU), named after G.M.B. Dobson, who investigated the ozone between 1920 and 1960. One Dobson Unit (DU) is defined to be 0.01 mm thickness of ozone at STP when all the ozone in the air column above an area is collected and spread over the entire area. Thus, 100 DU is 1 mm thick.

What is UV?

In the electromagnetic radiation spectrum, the region beyond the violet (wavelength ~ 400 nanometer nm) invisible to eye detection is called ultraviolet (UV) rays. Its wavelength is shorter than 400 nm.



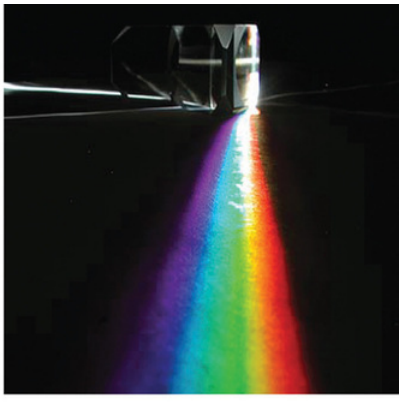


Figure 2.2.3.4.1: (a) This diagram shows the wavelength and frequency ranges of electromagnetic radiation. The visible portion of the electromagnetic spectrum is the narrow region with wavelengths between about 400 and 700 nm. (b) When white light is passed through a prism, it is split into light of different wavelengths, whose colors correspond to the visible spectrum.

UV is divided into three regions:

- UV A, wavelength = 400 - 320 nm
- UV B, wavelength = 320 - 280 nm
- UV C, wavelength = < 280 nm

Obviously, photons of UV C are the most energetic. UV-A radiation is needed by humans for the synthesis of vitamin-D; however, too much UV-A causes photoaging (toughening of the skin), suppression of the immune system and, to a lesser degree, reddening of the skin, and cataract formation. Ozone strongly absorbs UV B and C, but the absorption decreases as the wavelength increases to 320 nm. Very little UV C reaches the Earth surface due to ozone absorption.

How is Ozone produced in the Atmosphere?

When an oxygen molecule receive a photon (h\nu), it dissociates into monoatomic (reactive) atoms. These atoms attack an oxygen molecule to form ozone, O3.

$$\mathrm{O_2 + h\nu \rightarrow O + O} \tag{2.2.3.4.1}$$

$$\mathrm{O_2 + O \ \rightarrow O_3} \tag{2.2.3.4.2}$$

The last reaction requires a third molecule to take away the energy associated with the free radical $O^{\cdot}$ and O_2, and the reaction can be represented by

$$\mathrm{O_2 + O + M \ \rightarrow O_3 + M\cdot} \tag{2.2.3.4.3}$$

The over all reaction between oxygen and ozone formation is:

$$\mathrm{3\,O_2 \ \rightleftharpoons 2\,O_3} \tag{2.2.3.4.4}$$

The absorption of UV B and C leads to the destruction of ozone

$$\mathrm{O_3 + h\nu \rightarrow O + O_2} \tag{2.2.3.4.5}$$

$$\mathrm{O_3 + O \ \rightarrow 2\,O_2} \tag{2.2.3.4.6}$$

A dynamic equilibrium is established in these reactions. The ozone concentration varies due to the amount of radiation received from the sun.

2.2.3.4.1

The enthalpy of formation of ozone is 142.7 kJ / mol. The bond energy of O2 is 498 kJ / mol. What is the average O=O bond energy of the bent ozone molecule O=O=O?

Solution

The overall reaction is

$$\mathrm{3\,O_2 \ \rightarrow 2\,O_3} \quad \Delta H = 286kJ \tag{2.2.3.4.7}$$

Note that 3 O=O bonds of oxygen are broken, and 4 O-O bonds of ozone are formed. If the bond energy of ozone is *E*, then

$$\begin{aligned} E &= (3 * 498 + 286)kJ/4mol \\ &= 445kJ/mol \end{aligned}$$

DISCUSSION
The ozone bonds are slightly weaker than the oxygen bonds. The average bond energy is not the bond energy for the removal of one oxygen from ozone.

$$\mathrm{O_3 + h\nu \rightarrow O + O_2} \tag{2.2.3.4.8}$$

Can the energy to remove one oxygen be estimated from the data given here?

The techniques used in this calculation is based on the **principle of conservation of energy**.

Example 2.2.3.4.2

The bond energy of O2 is 498 kJ / mol. What is the maximum wavelength of the photon that has enough energy to break the O=O bond of oxygen?

Solution

The energy per O=O bond is:

$$(498000J/mol)/(6.022x1023bonds/mol) = 8.27x10 - 19J/bonds \tag{2.2.3.4.9}$$

The wavelength λ of the photons can be evaluated using

$$E = \frac{hc}{\lambda} \tag{2.2.3.4.10}$$

$$\lambda = \frac{(6.626 \times 10^{-34}\, J \cdot s) * (3 \times 10^{8}\, m/s)}{8.27 \times 10^{-19} J}$$
$$= 2.403 \times 10^{-7}\, m = 240 nm$$

DISCUSSION

The visible region range from 300 nm to 700 nm, and radiation with a wavelength of 240 nm is in the ultraviolet region (Figure 2.2.3.4.1). Visible light cannot break the O=O bond, and UV light has enough energy to break the O=O bond.

Chlorofluorohydrocarbons (CFCs)

Chemist Roy J. Plunkett discovered tetrafluoroethylene resin while researching refrigerants at DuPont. Known by its trade name, Teflon, Plunkett's discovery was found to be extremely heat-tolerant and stick-resistant. After ten years of research, Teflon was introduced in 1949. His continued research led to the usage of chlorofluorohydrocarbons known as CFCs or freon as refrigerants.

CFCs are made up of carbon, hydrogen, fluorine, and chlorine. DuPont used a number system to distinguish their product based on three digits. The digits are related to the molecular formulas.

- The first digit is the number of carbon atoms minus 1.
- The second digit is the number of hydrogen atoms plus 1.
- The third digit is the number of fluorine atoms minus 1.

For example, CFC (or freon) 123 should have a formula $C_2HF_3Cl_2$. The number of chlorine atoms can be deduced from the structural formula of saturated carbon chains. CFC's containing only one carbon atom per molecule has only two digits. Freon 12 used for fridge and automobil air conditioners has a formula of CF_2Cl_2. The nontoxic and nonflammable CFCs have been widely used as refrigerants, in aerosol spray, and dry cleaning liquid, foam blowing agents, cleansers for electronic components in the 70s, 80s and early 90s.

In 1973, James Lovelock demonstrated that all the CFCs produced up to that time have not been destroyed, but spread globally throughout the troposphere. (Lovelock's report was later published: J. E. Lovelock, R.J.Maggs, and R.J. Wade, (1974); *Nature*, **241**, 194) In the article, concentrations of CFCs at some parts per 10^{11} by volume was measured, and they deducted that with such a concentration, CFCs are not destroyed over the years. In 1974, Mario J. Molina published an article in *Nature* describing the ozone depletion by CFCs. (see M.J. Malina and F.S. Rowland, (1974); *Nature*, **249**, 810) NASA later confirmed that HF was present in the stratosphere, and this compound had no natural source but from the decomposition of CFCs. Molina and Rowland suggested that the chlorine radicals in CFCs catalyze the decomposition of ozone as discussed below.

How do CFCs help depleting ozone?

A relatively recent concern is the depletion of ozone, O3 due to the presence of chlorine in the troposphere, and eventually their migration to the stratosphere. A major source of chlorine is Freons: CFCl3 (Freon 11), CF2Cl2 (Freon 12), C2F3Cl3 (Freon 113), C2F4Cl2 (Freon 114). Freons decompose in the troposphere. For example,

$$CFCl_3 \rightarrow CFCl_2 + Cl \tag{2.2.3.4.11}$$

$$CF_2Cl_3 \rightarrow CF_2Cl + Cl^{\bullet} \tag{2.2.3.4.12}$$

The chlorine atoms catalyze the decomposition of ozone,

$$Cl + O_3 \rightarrow ClO + O_2 \tag{2.2.3.4.13}$$

and ClO molecules further react with O generated due to photochemical decomposition of ozone:

$$O_3 + h\nu \rightarrow O + O_2 \tag{2.2.3.4.14}$$

$$ClO + O \rightarrow Cl + O_2 \tag{2.2.3.4.15}$$

$$O + O_3 \rightarrow O_2 + O_2 \cdot \tag{2.2.3.4.16}$$

The net result or reaction is

$$2\,O_3 \rightarrow 3\,O_2 \tag{2.2.3.4.17}$$

Thus, the use of CFCs is now a world wide concern. In 1987, one hundred and forty nine (149) nations signed the Montreal Protocol. They agreed to reduce the manufacturing of CFCs by half in 1998; they also agree to phase out CFCs.

Ozone depletion in the polar region is different from other regions. The debate of ozone depletion often involves the North and South Poles. In these regions when temperatures drop to 190 K, ice cloud is formed. The ice crystals act as heterogeneous catalyst converting HCl and $ClONO_2$ into HNO_3 and Cl_2,

$$Cl + ClONO_2 \rightarrow HNO_3 + Cl_2 \tag{2.2.3.4.18}$$

$$H_2O + ClONO_2 \rightarrow HNO_3 + HOCl \cdot \tag{2.2.3.4.19}$$

Both Cl_2, and HOCl are easily photolyzed to Cl atoms, which catalyze the depletion of ozone. This has just been discussed in the previous section.

What has been done and what can be done to reduce ozone depletion?

The U.S. and Canadian governments have banned the use of Freons in aerosol sprays, but their use in air conditioner and cooling machines continue. In order to eliminate Freon in the atmosphere, international concerted effort and determination is required. However, sound and reliable scientific information is required. The banning of CFCs opens a research opportunity for another invention to find its substitute. Who knows what other problems will the new product bring?

Questions

1. **What is the unit used for measuring ozone layers?**

   Skill - Define a unit you use.

2. **What is the wavelength range of the UV radiation?**

   Skill - Describe UV radiation.

3. **How is ozone different from oxygen?**

   Skill - Describe the formation of ozone.

4. **When CFCs are exposed to UV or sun light, what species are produced?**

   Skill - Explain a photodecomposition reaction.

5. **What is the role of chlorine radical in the ozone formation or reactions.**

   Skill - Explain the mechanism of the catalytic reaction.

6. **What in the polar zone makes the depletion of ozone more serious?**

Solutions

1. Dobson unit.
2. UV radiation is electromagnetic radiation with wavelength between 100 and 400 nm.
3. The ozone molecules consist of 3 atoms whereas the usual oxygen molecules 2.
4. Reactive radicals are produced, including monoatomic chlorine radical.
5. Chlorine radical catalyze the decomposition of ozone.
6. The ice clouds act as heterogeneous catalysts for the formation of chlorine gas.
   Discussion - The chlorine gas is photodissociated into Cl radicals that catalyze the decomposition of ozone.

Learning Guide

- Arrange the regions of the electromagnetic spectrum in increasing energies of their photons: X-rays, visible, gamma rays, ultraviolet, infrared, microwave, etc.
- Examples on this *page* can be testing questions.

- What are CFCs?
  What are freon 12, 123, and 114?
  Explain how CFC help destroy the atmosphere ozone layer, particularly the polar region?

Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

This page titled 2.2.3.4: Ozone is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Chung (Peter) Chieh.

- **Ozone** by Chung (Peter) Chieh is licensed CC BY-NC-SA 4.0.

2.2.3.5: Photosynthesis

Discussion Questions

- How is light energy harvested in photosynthesis and what is the reaction center?

Photosynthesis

Photosynthesis is the process of converting light energy ($E = h\,\nu$) to chemical energy and storing it in the chemical bonds of sugar-like molecules. This process occurs in plants and some algae (Kingdom Protista). Plants need only light energy, CO_2, and H_2O to make sugar. The process of photosynthesis takes place in the chloroplasts (chloro = green; plasti = formed, molded), specifically using chlorophyll (phyll = leaf), the green pigment involved in photosynthesis.

How is light energy harvested in photosynthesis and what is the reaction center?

As early as 1640, people have demonstrated that photosynthesis was a means of converting carbon dioxide in the air into plant material. Thus, photosynthesis has been studied by many scientists, and there are many sites related to photosynthesis. Some of these sites describe their recent recent research activities, whereas others are for educational purposes. Even books and journals are introduced in this link.

The photosynthesis is a very complicated process, and we can only give an introduction here. Photosynthesis reduces carbon dioxide into carbohydrates,

$$6\,CO_2 + 12\,H_2O + hv \longrightarrow C_6H_{12}O_6 + 6\,O_2 + 6\,H_2O \qquad (2.2.3.5.1)$$

Thus, both electrons and energy are required. The electrons come from water molecules, and the energy is first absorbed by pigments known as chlorophylls and carotenoids. The former absorb blue (wavelength 430 nm) and red (wavelength 670 nm) light and the later absorbe blue-green light (wavelenghts between 400 and 500 nm). Green and yellow light are not absorbed. Reflection of these types of light makes plants appear green.

There are many varieties of pigments. They are bonded to proteins which provide pigment molecules with the appropriate orientation and position with respect to each other. After absorption by pigment, light energy is transferred to chlorophylls that are bonded to special proteins. Pigments and protein involved with this actual primary electron transfer event together are called the **reaction center**. A large number of pigment molecules (100-5000), collectively referred to as **antenna**, "harvest" light and transfer the light energy to the same reaction center. The purpose is to maintain a high rate of electron transfer in the reaction center, even at lower light intensities.

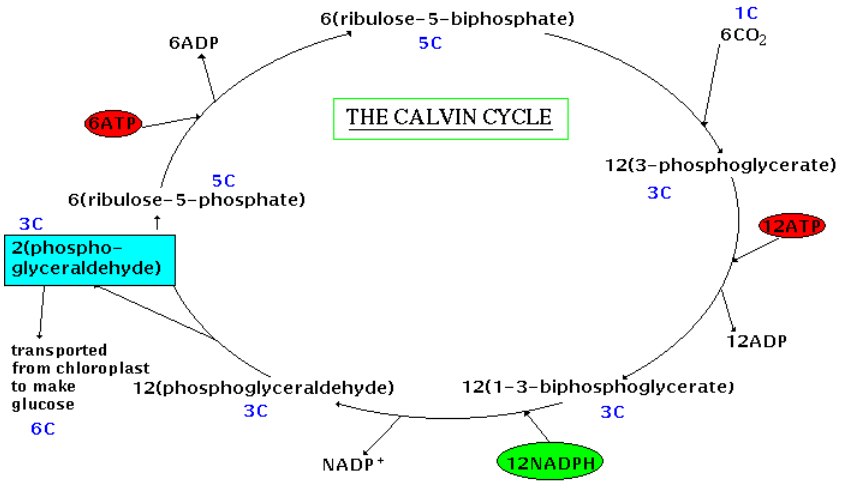
What are produced and consumed in plants respiration?

Photosynthesis is responsible for the production of oxygen and carbohydrates in plants. All living organisms respire, and so do plants. In the respiration, oxygen is consumed and carbon dioxide is produced.

respiration takes place all the time, but respiration is masked by higher rate of photosynthesis when the light intensities is high.

What is the Calvin-Benson cycle?

Working with the green algae chlorella, Melvin Calvin and Andy Benson, at the University of California at Berkeley, elucidated the following pathway for the conversion of carbon dioxide into carbohydrates:



Contributors and Attributions

- Chung (Peter) Chieh (Professor Emeritus, Chemistry @ University of Waterloo)

This page titled 2.2.3.5: Photosynthesis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Chung (Peter) Chieh.

- **Photosynthesis** by Chung (Peter) Chieh is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

2.2.4: Biochemical Cycles

2.2.4.1: Biochemical Cycles

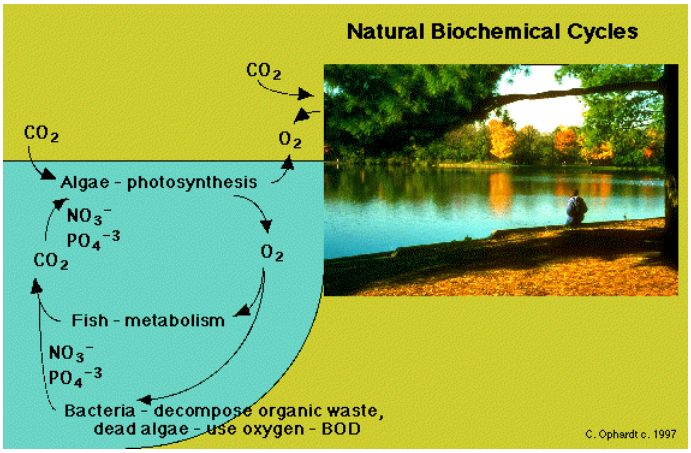
2.2.4.2: Carbon Cycle

2.2.4.3: Nitrogen Cycle

2.2.4: Biochemical Cycles is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2.4.1: Biochemical Cycles

Plants such as trees and algae undergo the photosynthesis reaction where carbon dioxide and water in the presence of sunlight are converted to organic materials and oxygen. An important reverse reaction occurs in the water: Fish use metabolism where oxygen and organic materials - other small fish or algae - as food is converted to carbon dioxide, water, and energy. Bacteria in water, as well as land, also undergo metabolism and use oxygen and decompose organic wastes as food to convert to carbon dioxide, water, and energy. By products in the decomposition of organic waste are nitrates and phosphates. The major natural biochemical cycles include the carbon, nitrogen, and phosphate cycles. They are presented in brief in this graphic.



The overall health of a body of water depends upon whether these factors are in balance. Municipal sewage systems are now doing a better job of removing most of the organic waste products in the discharge water, but some organic waste still enters the streams and lakes. If an excess amount of organic waste is present in the water, the bacteria use all of the available oxygen in the water in an attempt to decompose the organic waste.

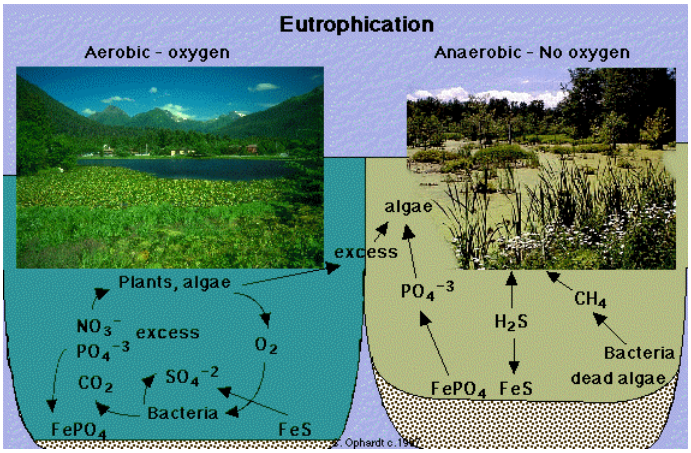
The amount of organic waste in water is represent by a chemical test called BOD - Biological Oxygen Demand. The concentration of oxygen is measured in a water sample at the beginning of the test and again after five days. The difference between the oxygen concentrations represents the amount of oxygen consumed by the bacteria in the metabolism of the waste organics present.

Eutrophication

In situations where eutrophication occurs, the natural cycles are overwhelmed by an excess of one or more of the following: nutrients such as nitrate or phosphate, or organic waste.

In the first case under aerobic conditions (presence of oxygen), the natural cycles may be more or less in balance until an excess of nitrate and/or phosphate enters the water. At this time the water plants and algae begin to grow more rapidly than normal. As this happens there is also an excess die off of the plants and algae as sunlight is blocked at lower levels. Bacteria try to decompose the organic waste, consuming the oxygen, and releasing more phosphate and nitrate to begin the cycle anew. Some of the phosphate may be precipitated as iron phosphate to remove the soluble form from the water solution.

In the second case under anaerobic conditions (absence of oxygen), as conditions worsen as more phosphates and nitrates may be added to the water, all of the oxygen may be used up by bacteria in trying to decompose all of the waste. Different bacteria continue to carry on decomposition reactions, however the products are drastically different. The carbon is converted to methane gas instead of carbon dioxide, sulfur is converted to hydrogen sulfide gas. Some of the sulfide may be precipitated as iron sulfide. Under anaerobic conditions the iron phosphate in the sediments may be solubilized into solution to make it available as a nutrient for the algae which would start the growth and decay cycle over again. The pond may gradually fill with undecayed plant materials to make a swamp.

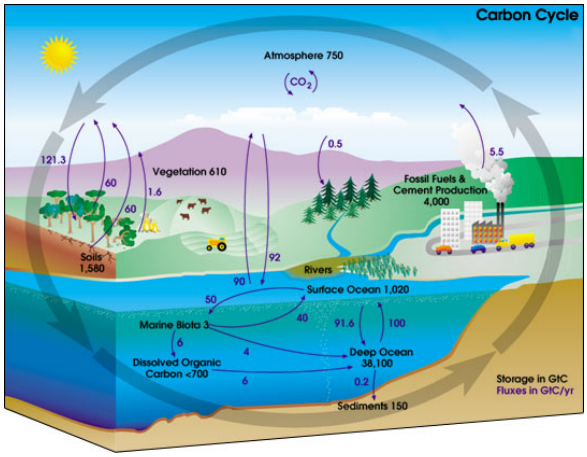


Contributors

- Charles Ophardt, Professor Emeritus, Elmhurst College; Virtual Chembook

2.2.4.1: Biochemical Cycles is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2.4.2: Carbon Cycle



Photosynthesis

Photosynthesis is a complex series of reactions carried out by algae, phytoplankton, and the leaves in plants, which utilize the energy from the sun. The simplified version of this chemical reaction is to utilize carbon dioxide molecules from the air and water molecules and the energy from the sun to produce a simple sugar such as glucose and oxygen molecules as a by product. The simple sugars are then converted into other molecules such as starch, fats, proteins, enzymes, and DNA/RNA i.e. all of the other molecules in living plants. All of the "matter/stuff" of a plant ultimately is produced as a result of this photosynthesis reaction. An important summary statement is that during photosynthesis plants use carbon dioxide and produce oxygen.

Carbon Cycle - Photosynthesis

$CO_2 + H_2O + energy \rightarrow C_6H_{12}O_6 + O_2$

carbon dioxide; water; sun; glucose; oxygen

represents any organic matter in plants

Combustion/Metabolism Reaction

Combustion occurs when any organic material is reacted (burned) in the presence of oxygen to give off the products of carbon dioxide and water and ENERGY. The organic material can be any fossil fuel such as natural gas (methane), oil, or coal. Other organic materials that combust are wood, paper, plastics, and cloth. Organic materials contain at least carbon and hydrogen and may include oxygen. If other elements are present they also ultimately combine with oxygen to form a variety of pollutant molecules such as sulfur oxides and nitrogen oxides.

Metabolism occurs in animals and humans after the ingestion of organic plant or animal foods. In the cells a series of complex reactions occurs with oxygen to convert for example glucose sugar into the products of carbon dioxide and water and ENERGY. This reaction is also carried out by bacteria in the decomposition/decay of waste materials on land and in the water.

An important summary statement is that during combustion/metabolism oxygen is used and carbon dioxide is a product. The whole purpose of both processes is to convert chemical energy into other forms of energy such as heat.

Carbon Cycle - Combustion, Metabolism

CH_4 or $C_6H_{12}O_6$ + O_2 → $CO_2 + H_2O$ + energy

oxygen; carbon dioxide; water; (heat)

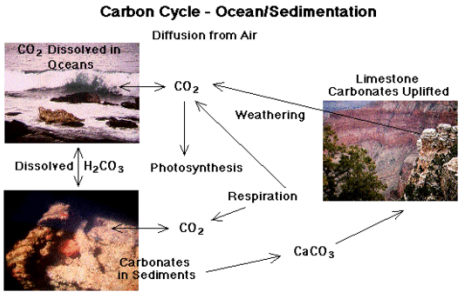
represents any fossil fuel, food, organic matter

Sedimentation

Carbon dioxide is slightly soluble and is absorbed into bodies of water such as the ocean and lakes. It is not overly soluble as evidenced by what happens when a can of carbonated soda such as Coke is opened. Some of the dissolved carbon dioxide remains in the water, the warmer the water the less carbon dioxide remains in the water.

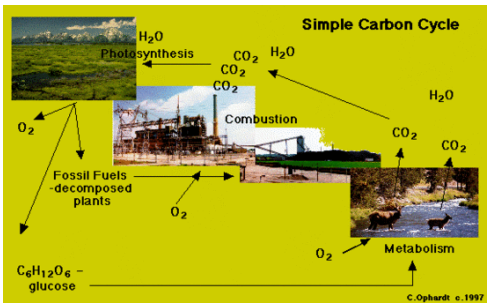
Some carbon dioxide is used by algae and phytoplankton through the process of photosynthesis.

In other marine ecosystems, some organisms such as coral and those with shells take up carbon dioxide from the water and convert it into calcium carbonate. As the shelled organisms die, bits and pieces of the shells fall to the bottom of the oceans and accumulate as sediments. The carbonate sediments are constantly being formed and redissolved in the depths of the oceans. Over long periods of time, the sediments may be raised up as dry land or into mountains. This type of sedimentary rock is called limestone. The carbonates can redissolve releasing carbon dioxide back to the air or water.



Human Impacts: Fossil Fuels

In the natural carbon cycle, there are two main processes which occur: photosynthesis and metabolism. During photosynthesis, plants use carbon dioxide and produce oxyge and during metabolism oxygen is used and carbon dioxide is a product. Humans impact the carbon cycle during the combustion of any type of fossil fuel, which may include oil, coal, or natural gas. Fossil Fuels were formed very long ago from plant or animal remains that were buried, compressed, and transformed into oil, coal, or natural gas. The carbon is said to be "fixed" in place and is essentially locked out of the natural carbon cycle. Humans intervene during by burning the fossil fuels. During combustion in the presence of air (oxygen), carbon dioxide and water molecules are released into the atmosphere. The question becomes as to what happens to this extra carbon dioxide that is released into the atmosphere. This is the subject of considerable debate and about it possible effect in enhancing the greenhouse ffect which may than result in global warming.



Contributors

- Charles Ophardt, Professor Emeritus, Elmhurst College; Virtual Chembook

2.2.4.2: Carbon Cycle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

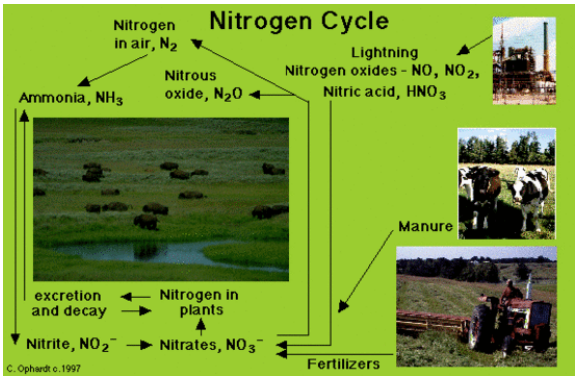
2.2.4.3: Nitrogen Cycle

The main component of the nitrogen cycle starts with the element nitrogen in the air. Two nitrogen oxides are found in the air as a result of interactions with oxygen. Nitrogen will only react with oxygen in the presence of high temperatures and pressures found near lightning bolts and in combustion reactions in power plants or internal combustion engines. Nitric oxide, NO, and nitrogen dioxide, NO_2, are formed under these conditions. Eventually nitrogen dioxide may react with water in rain to form nitric acid, HNO_3. The nitrates thus formed may be utilized by plants as a nutrient.

Nitrogen in the air becomes a part of biological matter mostly through the actions of bacteria and algae in a process known as nitrogen fixation. Legume plants such as clover, alfalfa, and soybeans form nodules on the roots where nitrogen fixing bacteria take nitrogen from the air and convert it into ammonia, NH_3. The ammonia is further converted by other bacteria first into nitrite ions, NO^{2-}, and then into nitrate ions, NO^{3-}. Plants utilize the nitrate ions as a nutrient or fertilizer for growth. Nitrogen is incorporate in many amino acids which are further reacted to make proteins.

Ammonia is also made through a synthetic process called the Haber Process. Nitrogen and hydrogen are reacted under great pressure and temperature in the presence of a catalyst to make ammonia. Ammonia may be directly applied to farm fields as fertilizer. Ammonia may be further processed with oxygen to make nitric acid. The reaction of ammonia and nitric acid produces ammonium nitrate which may then be used as a fertilizer. Animal wastes when decomposed also return to the earth as nitrates.

To complete the cycle other bacteria in the soil carry out a process known as denitrification which converts nitrates back to nitrogen gas. A side product of this reaction is the production of a gas known as nitrous oxide, N_2O. Nitrous oxide, also known as "laughing gas" - mild anesthetic, is also a greenhouse gas which contributes to global warming



Contributors

- Charles Ophardt, Professor Emeritus, Elmhurst College; Virtual Chembook

2.2.4.3: Nitrogen Cycle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

00: Front Matter

This page was auto-generated because a user created a sub-page to this page.

TABLE OF CONTENTS

00: Front Matter

- Table of Contents

2.3.1: Environmental Toxicology

- 2.3.1.1: Environmental toxicology
- 2.3.1.2: DPSIR
- 2.3.1.3: Short history

2.3.2: Environmental Chemistry, Chemicals

- 2.3.2.1: Introduction
- 2.3.2.2: Pollutants with specific properties
- 2.3.2.3: Pollutants with specific use

2.3.3: Environmental Chemistry - From Fate to Exposure

- 2.3.3.1: Environmental compartments
- 2.3.3.2: Sources of chemicals
- 2.3.3.3: Pathways and processes determining chemical fate
- 2.3.3.4: Partitioning and Partitioning Constants
- 2.3.3.5: Metal Speciation

2.3.4: Toxicology

- 2.3.4.1: Toxicokinetics
- 2.3.4.2: Toxicodynamics and Molecular Interactions
- 2.3.4.3: Toxicity Testing
- 2.3.4.4: Increasing ecological realism in toxicity testing
- 2.3.4.5: Availability and bioavailability
- 2.3.4.6: Degradation
- 2.3.4.7: Modelling exposure
- 2.3.4.8: Section 4.1. Toxicokinetics
- 2.3.4.9: Section 4.4. Increasing ecological realism in toxicity testing
- 2.3.4.10: Section 4.2. Toxicodynamics and Molecular Interactions
- 2.3.4.11: Section 4.3. Toxicity testing

2.3.5: Population, Community and Ecosystem Ecotoxicology

- 2.3.5.1: Introduction- Linking population, community and ecosystem responses
- 2.3.5.2: Population ecotoxicology in laboratory settings
- 2.3.5.3: Wildlife population ecotoxicology
- 2.3.5.4: Trait-based approaches
- 2.3.5.5: Population models
- 2.3.5.6: Metapopulations
- 2.3.5.7: Community ecotoxicology
- 2.3.5.8: Structure versus function incl. ecosystem services
- 2.3.5.9: Landscape ecotoxicology

2.3.6: Risk Assessment and Regulation

- 2.3.6.1: Introduction - The Essence of Risk Assessment
- 2.3.6.2: Ecosystem services and protection goals
- 2.3.6.3: Predictive risk assessment approaches and tools
- 2.3.6.4: Diagnostic risk assessment approaches and tools
- 2.3.6.5: Regulatory Frameworks
- 2.3.6.6: Risk management and risk communication
- 2.3.6.7: Risk perception

SECTION OVERVIEW

2.3.1: Environmental Toxicology

2.3.1.1: Environmental toxicology

2.3.1.2: DPSIR

2.3.1.3: Short history

This page titled 2.3.1: Environmental Toxicology is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

2.3.1.1: Environmental toxicology

1.1. Environmental toxicology

If you want to re-use this chapter, for e.g. in your electronic learning environment, feel free to copy this url: maken.wikiwijs.nl/120137/1__Introduction

Author: Ad Ragas

Reviewers: Kees van Gestel, Nico van Straalen

Learning objectives

You should be able to

- characterize the field of environmental toxicology;
- explain what type of knowledge from environmental chemistry, toxicology and ecology is relevant for environmental toxicology;
- explain the difference between environmental toxicology and ecotoxicology.

Keywords: Environmental toxicology, environmental chemistry, toxicology, ecology, ecotoxicology

Environmental toxicology is the science that studies the fate and effects of potentially hazardous chemicals in the environment. It is a multidisciplinary field assimilating and building upon knowledge, concepts and techniques from other disciplines, such as toxicology, analytical chemistry, biochemistry, genetics, ecology and pathology. Environmental toxicology emerged in response to the growing awareness in the second part of the 20^{th} century that chemicals emitted to the environment can trigger hazardous effects in organisms living in this environment, including humans. Section 1.3 gives a brief summary of the history of environmental toxicology.

One way to depict the field of environmental toxicology is by a triangle consisting of chemicals, the environment and organisms (Figure 1). The triangle illustrates that the fate and potential hazardous effects of chemicals emitted to the environment are determined by the interactions between these chemicals, the environment and organisms. The fate of substances in the environment is the topic of **environmental chemistry**, the effects of substances on living organisms is studied by **toxicology**, and the implications of these effects on higher levels of biological organization are analyzed by the field of **ecology**.

Another term widely used to refer to this field of study is **ecotoxicology**. The main distinction is the inclusion of human health as an endpoint in environmental toxicology, whereas ecotoxicology is restricted to ecological endpoints. Since the current book includes human health as an assessment endpoint for environmental contaminants, the term environmental toxicology is preferred over ecotoxicology.

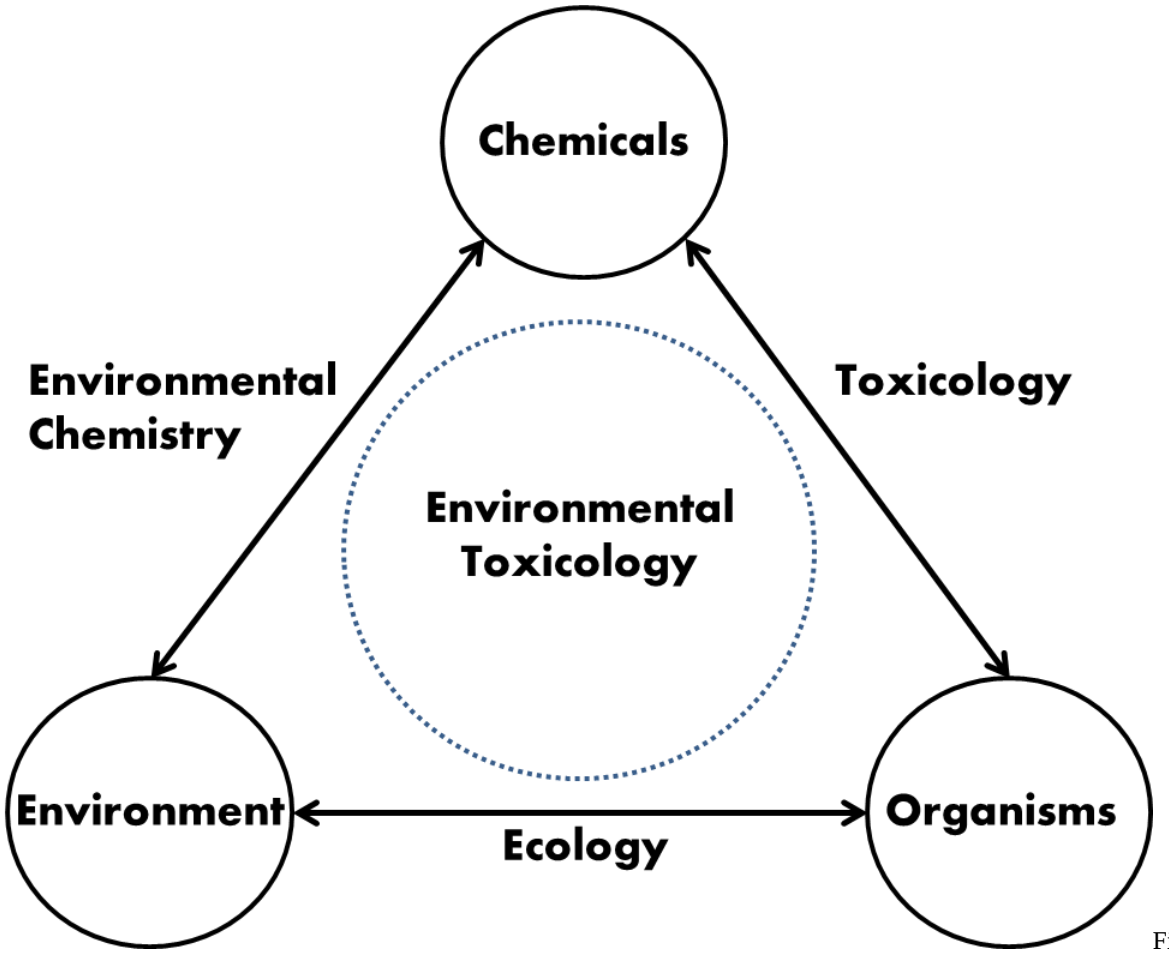


Figure 1: Environmental toxicology studies the interactions between chemicals, organisms and the environment making use of environmental chemistry, toxicology and ecology. Source: Ad Ragas.

Environmental chemists study the fate of chemicals in the environment, e.g. their distribution over different environmental compartments and how this distribution is influenced by the physicochemical properties of a chemical and the characteristics of the environment. They aim to understand the pathways and processes involved in the environmental fate of a chemical after it has been emitted to the environment, including processes such as advection, deposition and (bio)degradation. Within the context of environmental toxicology, the ultimate aim is to produce a reliable assessment of the exposure of organisms, an aim which is often complicated by the enormous heterogeneity of the environment.

Environmental chemists use a variety of tools to analyze and assess the fate of chemicals in the environment. Two fundamental tools are analytical measurements and mathematical modelling. Measurements are essential to acquire new knowledge and insight into the behavior of chemicals in the environment., e.g. measurements on emissions, environmental concentrations and specific processes such as biodegradation. These measurements are analyzed to discover patterns, e.g. between substance properties and environmental characteristics. Once revealed, such patterns can be integrated into a comprehensive mathematical model to predict the fate of and exposure to substances in the environment. If sufficiently validated, these models can subsequently be used by risk assessors to assess the exposure of organisms to chemicals, reducing the need for expensive measurements.

Chapter 2 focuses on the types of chemicals occurring in the environment, their sources and the concentrations found at contaminated sites. In Chapter 3, focus will be on the fate and transport of these chemicals, including aspects of bioavailability and bioaccumulation in organisms.

Toxicologists study the effects of chemicals on organisms, often at the individual level. Fundamental toxicologists aim to understand the mechanisms involved in the toxicity of a compound, whereas more applied toxicologists are primarily interested in the relationship between exposure and effect, often with the aim of identifying an exposure level that can be considered safe.

Within this context, the dose concept as introduced by Parcelsus at the start of the 16th century is essential (see Section 1.3), i.e. the likelihood of adverse effects depends on the dose organisms are being exposed to.

The processes taking place after exposure of an organism to a toxicant are often divided into toxicokinetic and toxicodynamic processes. Toxicokinetic processes are those that describe the fate of the toxicant in the organism, including processes such as absorption, distribution, metabolism and excretion (ADME). These toxicokinetic or ADME processes are sometimes collectively referred to as "What the body does to the substance" and determine the exposure level at the site of toxic action, or internal dose. Toxicodynamic processes are those that describe the evolution of an adverse effect from the moment that the toxicant, or one of its metabolites, interacts with a molecular receptor in the body. This interaction is often referred to as the primary lesion or molecular initiating event (MIE). Toxicodynamic processes are sometimes collectively referred to as "What the substance does to the body" and the chain of events leading to an adverse outcome as the adverse outcome pathway (AOP).

The toxicity of a compound thus depends on toxicokinetic as well as toxicodynamic processes. Traditionally, this toxicity is being determined by exposing whole organisms in the laboratory to the substance of interest, and subsequently monitoring the health status of these organisms. However, as a result of the growing societal pressure to reduce animal testing, as well as the increased mechanistic understanding and improved molecular techniques, this so-called "black box approach" is more and more being replaced by a combination of in vitro toxicity testing and "in silico" predictive approaches. Physiologically-based toxicokinetic (PBTK) models are increasingly used to model the fate of chemicals in the body, resulting in a prediction of the internal exposure. *In vitro* tests and advanced molecular techniques at the gene (genomics) or protein (proteomics) level may subsequently be used to determine whether these internal exposure levels will trigger adverse effects, although many challenges remain in the prediction of adverse effects based on *in vitro* test and omics information. Chapter 4 focuses on dose-response relationships, modes of action, species differences in sensitivity and resistance against toxicants.

Ecologists study the interactions between organisms and their environment. Ecology is an important pillar of environmental toxicology, because ecological knowledge is needed to translate effects at the individual level to the ecosystem level; an important endpoint of ecological risk assessments. Such a translation requires specific knowledge, e.g. on life cycles of organisms, natural factors regulating their populations, genetic variability within populations, spatial distribution patterns, and the role organisms play in processes like nutrient cycling and decomposition. Effects considered relevant at the individual level, such as a tumor risk, may turn out to be irrelevant at the population or ecosystem level. Similarly, subtle effects at the individual level may turn out to be highly relevant at the ecosystem level, e.g. behavioral changes after environmental exposure to antidepressants which may affect the population dynamics of fish species. In recent years, there is an increasing interest for the role of the landscape configuration, distribution patterns and their dynamics in environmental toxicology. The spatial configuration of the landscape, the distribution of species and the timing of exposure events turn out to be important determinants of ecosystem effects. The ecological aspects of environmental toxicology will be discussed in Chapter 5.

1.1. Question 1

What is the difference between environmental toxicology and ecotoxicology?

1.1. Question 2

Indicate from the following terms whether these belong to the environmental chemistry, toxicology or ecology?

(Bio)degradation

Fate and exposure model

Dose

Toxicodynamics

Adverse outcome pathway

Population dynamics

Landscape configuration

1.1. Question 3

Give an example how subtle effects which may remain undetected in a toxicity test can be relevant at the population or ecosystem level.

This page titled 2.3.1.1: Environmental toxicology is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **1.1: Environmental toxicology** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.1.2: DPSIR

1.2. DPSIR

Author: Ad Ragas

Reviewers: Frank van Belleghem

Learning objectives

You should be able to:

- list and describe the five categories of DPSIR;
- structure a simple environmental problem using the DPSIR framework;
- describe the position and role of environmental toxicology within the DPSIR framework;
- indicate the most important advantages and disadvantages of the DPSIR framework.

Keywords: Drivers, pressures, state variables, impacts, responses

On the one hand, environmental toxicology is rooted in more fundamental scientific disciplines like biology and chemistry where curiosity is an important driver for gathering new knowledge. On the other hand, environmental toxicology is a problem-oriented discipline. As such, it is part of the broader field of environmental sciences which analyses the interactions between society and its physical environment in order to promote sustainability. Within this context, knowledge about the interactions of substances with the biotic and abiotic environment is being generated with the ultimate aim to prevent and address potential pollution problems in society. To be able to contribute optimally, an environmental toxicologist should know how pollution problems are structured and what the role of environmental toxicologists is in analysing, preventing and solving such problems. A widely used framework for structuring environmental problems is DPSIR. DPSIR stands for Drivers, Pressures, State, Impacts and Responses (Figure 1). The aim of the current section is to explain the DPSIR framework.

Communication tool

Communication is essential when analysing and addressing societal issues such as environmental pollution. As an environmental toxicologist, you will have to communicate with fellow scientists to develop a common understanding of the pollution problem, and with policy makers and stakeholders (e.g., producers of chemicals and consumers that are being exposed to chemicals) to explain the scientific state of the art. It is likely that you will use terms like "cause", "source" and "effects". However, not everybody will use and perceive these terms in the same way. Some people may argue that a farmer is the main cause of pesticide pollution, whereas others may argue that it is the pesticide manufacturer, or even the increasing world population. Likewise, some people may perceive the concentration of pesticides in water as an effect of pesticide use, whereas others may refer to the extinction of species when talking about effects. These differences may result in miscommunication, complicating scientific analysis and the search for appropriate solutions.

The DPSIR framework is a tool that helps preventing such communication problems. It provides a common and flexible frame of reference to structure environmental issues by describing these in terms of drivers, pressures, state (variables), impacts and responses (Figure 1). Flexibility is an important characteristic of the framework, enabling adaptation to the problem at hand. The DPSIR framework should not be considered a panacea or used as a mould that rigidly fits all environmental issues. Its main strength is that it stimulates communication between scientists, policy makers and other actors and thereby supports the development of a common understanding.

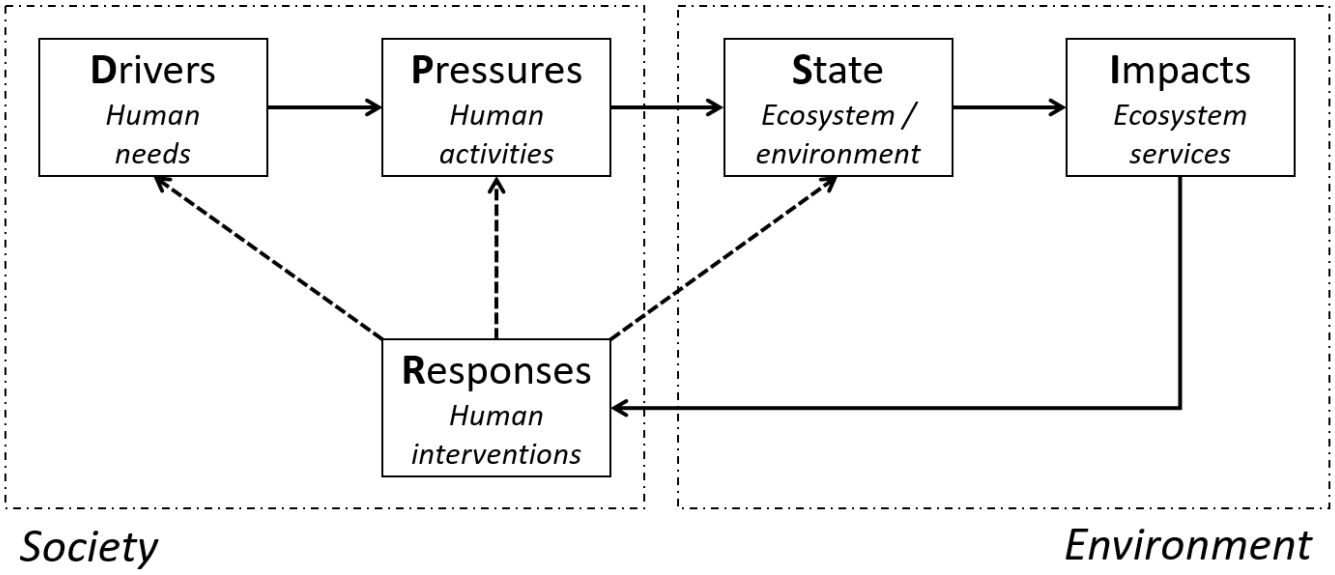


Figure 1. The DPSIR framework is a tool to structure environmental issues by organizing the processes in Drivers, Pressures, State (variables), Impacts and Responses. Source: Ad Ragas.

The framework

The DPSIR framework essentially is a cause-and-effect chain that aims to capture the main processes involved in an environmental issue; from its origin to the changes it triggers in the environment and in society. These processes are organized in five main categories, i.e.:

- are the human needs underlying the human activities that ultimately result in adverse effects. An example is the human need for food resulting in the use of pesticides such as neonicotinoids.
- are human activities initiated to fulfil human needs and resulting in changes in the physical environment that ultimately lead to - often unforeseen - adverse consequences for the environment or certain groups of society that are perceived as problematic, either now or in the future. An example is the use of neonicotinoids in agriculture.
- refers to the status of the physical environment. The state of the environment is often quantified using observable changes in environment parameters, e.g., the concentration of neonicotinoids in water, air, soil and biota.
- are any changes in the physical environment or society that are a consequence of the environmental pressures and that are perceived as problematic by society or some groups in society. An example is the increasing bee mortality that is, at least partly, attributed to the use of neonicotinoids. Or the human health effects of pesticides.
- are all initiatives developed by society to address the issue. These can range from gathering knowledge to developing policy plans and taking measures to mitigate effects or reduce emissions. Examples include the introduction of a risk-based admission procedure for neonicotinoids, the introduction of more efficient spraying techniques, and the development of environmentally friendly pest control techniques.

In principle, any environmental issue can be captured in a DPSIR. But it is important to realize that the labelling of processes as either drivers, pressures, state (variables), impacts or responses is likely to differ between people since the categories are broadly defined and the level of detail in the processes considered may vary. For example, some people may argue that "agriculture" should be classified as a driver, whereas others may argue it is a pressure. Yet other people may deal with this issue by adapting the DPSIR framework, i.e. by adding a new category called "human activities" that is placed in-between the drivers and the pressures. Another typical issue is the labelling of consecutive changes in the physical environment such as rising CO_2 levels, increases in temperature and changes in species abundance. These changes can be labelled as changes in consecutive state variables, i.e. state variables of 1^{st}, 2^{nd} and 3^{rd} order. The idea is that 1^{st} order changes trigger 2^{nd} order changes, e.g. rising CO_2 levels triggering a rise in temperature, and 2^{nd} order changes trigger 3^{rd} order changes, in this case a shift in species abundance. The change in species abundance may also be labelled as an impact, provided this change is perceived as problematic by (groups in) society. The category "impacts" is closely related to the protection goals of risk assessment (see the Section Ecosystem services and protection goals). If

there is consensus in society that an impact should be prevented, it becomes a protection goal. All these examples illustrate that the DPSIR framework should be applied in a flexible way and that communication is essential.

Environmental toxicology mainly focuses on the Pressures, State and Impacts blocks of the DPSIR chain. The use of chemicals by society, e.g. in agriculture or in consumer products, and their emission to the environment belongs to the Pressure block. The fate of chemicals in the environment and their accumulation in organisms belongs to the State block. And the adverse effects triggered in ecosystems and humans belong to the Impact block. An important step in risk assessment of chemicals (Chapter 6) is the derivation of safe exposure levels such as the Predicted No Effect Concentration (PNEC) for ecosystems or the Acceptable Daily Intake (ADI) for humans. In terms of DPSIR, this boils down to defining an acceptable impact level (e.g. a zero effect level or a 1 in a million tumor risk) and translating this into a corresponding state parameter (e.g. the chemical concentration in air or water). Fate modelling (Section on Modelling exposure) aims to predict soil, water, air and organisms (all State parameters) based on emission data (a Pressure parameter).

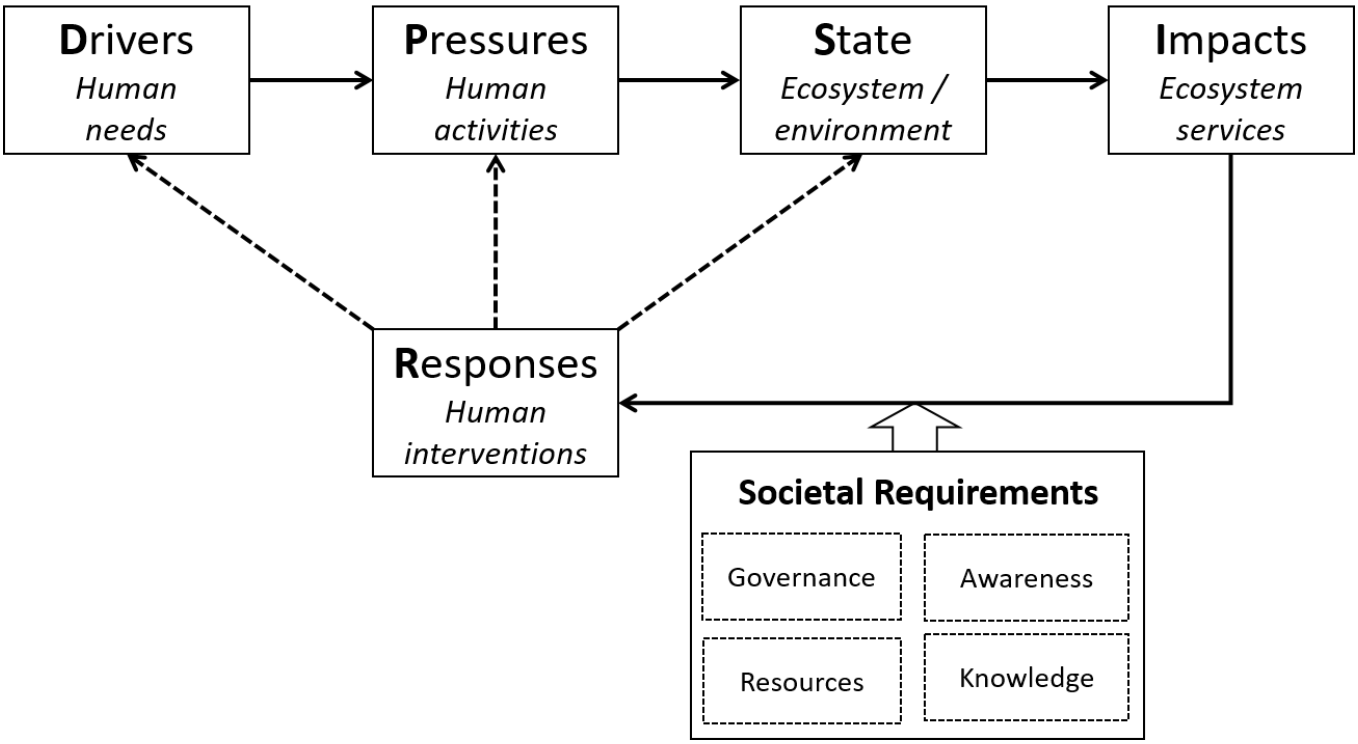


Figure 2. The extended DPSIR framework to put more emphasis on the societal dimension, i.e. governance, awareness, resources and knowledge. Source: Ad Ragas.

The DPSIR framework has been criticized because it tries to capture all processes in cause-and-effect relationships, resulting in a bias towards the physical dimension of environmental issues, e.g. human activities, emissions, physical effects and mitigations measures. The societal dimension is less easily captured, e.g. knowledge generation, governance structures, resources needed to implement measures, awareness and societal framing of the problem (Svarstad et al., 2008). Although the DPSIR framework can been adapted to accommodate some of these aspects (e.g., see Figure 2), it should be acknowledged that it has its limitations. Several alternative frameworks have been developed, and some of these better capture the societal dimension (Gari et al., 2015; Elliott et al., 2017). Nevertheless, DPSIR can be a useful framework to contextualize the problems that are addressed in environmental toxicology. It nicely shows why knowledge on the fate and impact of chemicals (state and impacts) is needed to address pollution issues and that the use of this knowledge is always subject to valuation, i.e. it depends on how society values the adverse effects triggered by the pollution. DPSIR is also widely used by national and international institutes such as the European Environment Agency (EEA), the United States Environmental Protection Agency (US-EPA) and the Organisation for Economic Cooperation and Development (OECD). The DPSIR framework is sometimes also used as a first step in modelling, especially its physical dimension. Once relevant processes have been identified, these are then described quantitatively resulting in models that can be used to predict environmental concentrations or ecological effects of substances based on knowledge about human activities or emissions.

References

Gari, S.R., Newton, A., Icely, J.D. (2015). A review of the application and evolution of the DPSIR framework with an emphasis on coastal social-ecological systems. Ocean & Coastal Management 103, 63-77.

Svarstad, H., Petersen, L.K., Rothman, D., Siepel, H., Wätzold, F. (2008). Discursive biases of the environmental research framework DPSIR. Land Use Policy 25, 116-125.

Elliott, M., Burdon, D., Atkins, J.P., Borja, A., Cormier, R., de Jonge, V.N., Turner, R.K. (2017). "And DPSIR begat DAPSI(W)R(M)!" - A unifying framework for marine environmental management. Marine Pollution Bulletin 118, 27-40.

1.2. Question 1

Indicate whether the following phenomena should be labelled as drivers, pressures, state (variables), impacts and responses.

1. The number of fish in a water body
2. The pesticide concentration in a water body
3. The development of a new spaying technique to reduce pesticide emissions
4. The need for food
5. Crop cultivation
6. Spraying pesticides

1.2. Question 2

Pharmaceuticals are being used to protect the health of humans, farm animals and pets. After use, part of these pharmaceuticals may reach the environment where they may trigger adverse effects in ecosystems. In theory, humans may also be affected, e.g. after swimming in polluted surface water or consumption of polluted drinking water. Besides direct toxic effects, antibiotics may also result in the emergence of antibiotic resistance, which threatens human health.

List the most important drivers, pressures, state (variables), impacts and responses for the issue of "pharmaceuticals in the environment".

1.2. Question 3

On which blocks of the DPSIR framework do you focus when you work as an environmental toxicologist?

This page titled 2.3.1.2: DPSIR is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **1.2: DPSIR** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.1.3: Short history

1.3. Short history

Author: Ansje Löhr

Reviewers: Ad Ragas, Kees van Gestel, Nico van Straalen

Learning Objective:

You should be able to

- summarize the history of environmental toxicology
- describe the increasing awareness over time of environmental and health risks

Keywords: Paracelsus; Rachel Carson (Silent Spring); Awareness; SETAC; standards

History

From earliest times, man has been confronted with the poisonous properties of certain plants and animals. Poisonous substances are indeed common in nature. People who still live in close contact with nature generally possess an extensive empirical knowledge of poisonous animals and plants. Poisons were, and still are, used by these people for a wide range of applications (catching fish, poisoning arrowheads, in magic rituals and medicines). The first Egyptian medical documentation (written in the Ebers Papyrus) dates from 1550 BC and demonstrates that the ancient Egyptians had an extensive knowledge of the toxic and curative properties of natural products. A good deal is known about the information regarding toxic substances possessed by the Greeks and the Romans. They were very interested in poisons and used them to carry out executions. Socrates, for example, was executed using an extract of hemlock (*Conium maculatum).* It was also not unusual to use a poison to murder political opponents. Poisons were ideal for that purpose, since it was usually impossible to establish the cause of death by examining the victim. To do so would have required advanced chemical analysis, which was not available at that time.

Early European literature also includes a considerable number of writings on toxins, including the so-called herbals, such as the Dutch "Herbarium of Kruidtboeck" by Petrus Nylandt dating from 1673. Poisoning sometimes assumed the character of a true environmental disaster. One example is poisoning by the fungus *Claviceps purpurea*, which occurs as a parasite in grain, particularly in rye (spurred rye) and causes the condition known as ergotism. In the past, this type of epidemic has killed thousands of people, who ingested the fungus with their bread. There are detailed accounts of such calamities. For example, in the year 992 an estimated 40,000 people died of ergotism in France and Spain. People were not aware of the fact that death was caused by eating contaminated bread. It was not until much later that it came to be understood that large-scale cultivation of grain involved this kind of risk.

Paracelsus

It was pointed out centuries ago that workers in the mining industry, who came into contact with a variety of metals and other elements, tended to develop specific diseases. The symptoms regularly observed as a result of contact with arsenic and mercury in the mining industry were described in detail by the famous Swiss physician Paracelsus (Figure 1) in his 1567 treatise "Von der Bergsucht und anderen Nergkrankheiten" (miners sickness and other diseases of mining). During the emergence of the scientific renaissance of the 16^{th} century, Paracelsus (1493 - 1541) drew attention to the dose-dependency of the toxic effect of substances. In the words of Paracelsus, "all Ding sind Gifft … allein die Dosis macht das ein Ding kein Gifft is" (everything is a poison … it is only the dose that makes it not a poison). This principle is just as valid today. At the same time, it is one of the most neglected principles in the public understanding of toxicology.

Figure 1: A portrait of Paracelsus (PORTRAIT PRESUME DU MEDECIN PARACELSE (1493-1541) (Source: https://commons.wikimedia.org/wiki/F...Paracelsus.jpg)

A work from the same period "De Re Metallica" by Gergius Agricola (Georg Bauer, 1556), deals with the health aspects of working with metals. Agricola even advised preventive aspects, such as wearing protective clothing (masks) and using ventilation.

Scrotum cancer in chimney sweepers: carcinogenicity of occupational exposure

Another example of the rising awareness of the effects of poisons on human health came with the suggestion, by Percival Pott in 1775, that the high frequency of scrotum cancer among British chimney sweepers was due to exposure to soot. He was the first to describe occupational cancer.

> A part of the essay by Percival Pott "The fate of these people seems singularly hard; in their early infancy, they are most frequently treated with great brutality, and almost starved with cold and hunger; they are thrust up narrow, and sometimes hot chimnies, where they are bruised, burned, and almost suffocated; and when they get to puberty, become peculiarly liable to a most noisome, painful, and fatal disease." See the rest of the original text of his essay here.

Soot consists of polycyclic aromatic hydrocarbons (PAHs) and their derivatives. The exposure to soot came with concurrent exposure to a number of carcinogens such as cadmium and chromium. From the 1487 cases of scrotal cancer reported, 6.9 % occurred in chimney sweepers. Scrotal and other skin cancers among chimney sweepers were at the same time also reported from several other countries.

Peppered moth in polluted areas

Changes in the environment due to environmental pollution led to interesting insights in the potential of species to adapt for survival and the role of natural selection in it. A famous example of such micro-evolution is the peppered moth, *Biston betularia*, that is generally a mottled light color with black speckles. This pattern gives them good camouflage against lichen-covered tree trunks while resting during the day. During the industrial revolution, the massive increase in the burning of coal resulted in the emission of dark smoke turning the light trees in the surrounding areas dark. As a consequence, the dark, melanic form of the peppered moth took over in industrial parts of the United Kingdom during the 1800s. The melanic forms used to be quite rare, but their dark color served as a protective camouflage from bird populations in the polluted areas. This allowed them to become dominant in areas with soot-covered trunks. Two British biologists, Cedric Clarke and Phillip Sheppard, discovered this when they pinned dead moths of the two types on dark and light backgrounds to study their predation by birds. The dark moths had an advantage in the dark forests, a result of natural selection. In areas where air pollution has decreased the melanic form became less abundant again.

Video on peppered moths

Awareness in the 1950s and 1960s

After the second world war, synthetic chemical production became widespread. However, there was limited awareness of the environmental and health risks. In the 1950s, Environmental Toxicology came to light as a result of increasing concern about the

impact of toxic chemicals on the environment. This led toxicology to expand from the study of the toxic impacts of chemicals on man to that of toxic impacts on the environment. An important person in raising this awareness was Rachel Carson. Her book "Silent Spring", published in 1962, in which she warned of the dangers of chemical pesticides, triggered widespread public concern on the dangers of improper pesticide use.

First have a look at an historical clip on the use of dichlorodiphenyltrichloroethane, commonly known as DDT, that was developed in the 1940s as the first modern synthetic insecticide.

https://www.youtube.com/embed/Ipbc-6IvMQI

Silent Spring - Rachel Carson

https://www.youtube.com/embed/SeJNRaE11A0

DDT is very persistent and tends to concentrate when moving through the food chain. As a consequence, the use of DDT led to very high levels, especially in organisms high in the food chain. Bioaccumulation in birds appeared to cause eggshell thinning and reproductive failure. Because of the increasing evidence of DDT's declining benefits and its environmental and toxicological effects, the United States Department of Agriculture, the federal agency responsible for regulating pesticide use, began regulatory actions in the late 1950s and 1960s to prohibit many of its uses. By the 1980s, the use of DDT was also banned from most Western countries.

Large environmental disasters

As a result of large environmental disasters, awareness amongst the general public increased. An enormous industrial pesticide disaster occurred in 1984 in Bhopal, India, when more than 40 ton of the highly toxic methyl isocyanate (MIC) gas leaked from a pesticide plant into the towns located near the plant. Almost 4000 people were killed immediately and 500,000 people were exposed to the poisonous substance causing many additional deaths because of gas-related diseases. The plant was actually initially only allowed to import MIC but was producing it on a large scale by the time of the disaster and safety procedures were far below (international) standards for environmental safety. The disaster made it very clear that this should be changed to avoid other large-scale industrial disasters.

The Sandoz agrochemical spill close to Basel in Switzerland in 1986 was the result of a fire in a storehouse. The emission of large amounts of pesticide caused severe ecological damage to the Rhine river and massive mortality of benthic organisms and fish, particularly eels and salmonids.

At the time of these incidents, environmental standards for chemicals were still largely lacking. The incidents triggered scientists to do more research on the adverse environmental impacts of chemicals. Public pressure to control chemical pollution increased and policy makers introduced instruments to better control the pollution, e.g. environmental permitting, discharge limits and environmental quality standards.

Our Common Future

In 1987, the World Commission on Environment and Development released the report "Our Common Future", also known as the Brundtland Report. This report placed environmental issues firmly on the political agenda, defining sustainable development as "a development that meets the needs of the present without compromising the ability of future generations to meet their own needs". Another influential book was "Our stolen future" written by Theo Colborn and colleagues in 1996. It raised awareness of the endocrine disrupting effects of chemicals released into the environment and threatening (human) reproduction by emphasizing it not only concerns feminization of fish or other organisms in the environment, but also the human species.

Please watch the video "Developments in Environmental Toxicology - Interview with two pioneers" included at the start of the Introduction of this book.

SETAC

Before the 1980s, no forum existed for interdisciplinary communication among environmental scientists -biologists, chemists, toxicologists- as well as managers and others interested in environmental issues. The Society of Environmental Toxicology and Chemistry (SETAC) was founded in North America in 1979 to fill this void. In 1991, the European branch started its activities and later SETAC also established branches in other geographical units, like South America, Africa and South-East Asia. SETAC publishes two journals: Environmental Toxicology and Chemistry (ET&C) and Integrated Environmental Assessment and Management (IEAM). SETAC also is active in providing training, e.g. a variety of online courses where you can acquire skills and insights in the latest developments in the field of environmental toxicology. Based on the growth in the society's membership, the meeting attendance and their publications, a forum like SETAC was clearly needed. Read more on SETAC, their publications and how you can get involved here.

Where SETAC focuses on environmental toxicology, international toxicological societies have also been established like EUROTOX in Europe and the Society of Toxicology (SOT) in North America. In addition to SETAC, EUROTOX and SOT, many national toxicological societies and ecotoxicological counterparts or branches became active since the 1970s, showing that environmental toxicology has become a mature field of science. One element indicative of this maturation, also is that the different societies have developed programmes for the certification of toxicologists.

References

Carson, R. (1962). *Silent Spring*. Houghton Mifflin Company.

Colborn, T., Dumanoski, D., Peterson Myers, J. (1996). *Our Stolen Future: Are We Threatening Our Fertility, Intelligence, and Survival? A Scientific Detective Story*. New York: Dutton. 306 p.

World Commission on Environment and Development (1987). *Our Common Future*. Oxford: Oxford University Press. p.27.

1.3. Question 1

Paracelsus is famous for the dose-dependency of the toxic effect of substances. What is meant by dose-dependency?

1.3. Question 2

What is the difference between toxicology and environmental toxicology?

1.3. Question 3

How was sustainable development defined in "Our Common Future" also known as the Brundtlandt report?

This page titled 2.3.1.3: Short history is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **1.3: Short history** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

SECTION OVERVIEW

2.3.2: Environmental Chemistry, Chemicals

2.3.2.1: Introduction

2.3.2.2: Pollutants with specific properties

2.3.2.3: Pollutants with specific use

This page titled 2.3.2: Environmental Chemistry, Chemicals is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

2.3.2.1: Introduction

2.1. Introduction

Authors: John Parsons, Steven Droge

Reviewer: Kees van Gestel

Leaning objectives:

You should be able to:

- mention the main groups of environmental pollutants
- comprehend the molecular structures of the most important organic pollutants
- mention the most important functional groups determining the environmental properties of organic pollutants

Keywords: natural toxicants, molecular structures, pollutant classes, anthropogenic pollutants

Introduction

Environmental toxicology deals with the negative effects of the exposure to chemicals we regard as pollutants (or contaminants/toxicants). Environmental toxicants receive a lot of media attention, but many critical details are getting lost or are easily forgotten. The clip "You, Me, DDT" shows the discovery of the grandson of the works of his grandfather, the Swiss inventor of the insecticide DDT, Paul Hermann Müller, who received the Nobel Prize for Medicine for that in 1948 (see also Section 1.3). The clip "Stop the POPs" interviews (seemingly) common people about one of the most heavily regulated group of pollutants.

Organisms, including humans, have always been exposed to chemicals in the environment and rely on many of these chemicals as nutrients. Volcanoes, flooding of acid sulfur lakes, and forest fires have caused widespread contamination episodes. Organisms are also in many cases directly or indirectly involved in the fate and distribution of undesirable chemicals in the environment. Many naturally occurring chemicals are toxicants already (see also Section 1.3), think for example about:

- local arsenic or mercury hotspots in the Earth's crust, contaminating water pumps or rice irrigation fields;
- plant-based defence chemicals such as alkaloids, morphine in poppy seeds, juglone from black walnut trees);
- fungal toxins, such as mycotoxins threatening grain storage depots after harvests;
- bacterial toxins, such as the botulinum toxin, a neurotoxic protein produced by the bacterium *Clostridium botulinum* which is the most acutely lethal toxin known at ~10 ng/kg body weight when inhaled;
- phycotoxins, produced by algae, in mass algal blooms or those that may end up at dangerous levels in shell food;
- zootoxins in animals, such as venom of snakes and defensive toxins on the skin of amphibians.

Human activities have had an enormous impact on the increased exposure to natural chemicals as a result of, for example, the mining and use of metals, salts and fossil fuels from geological resources. This is for example the case for many metals, nutrients such as nitrate, and organic chemicals present in fossil fuels. Additionally, the industrial synthesis and use of organic chemicals, and the disposal of wastes, have resulted in a wide variety of hazardous chemicals that had either never existed before, or at least not in the levels or chemical form that occur nowadays in our heavily polluted global system. These are typically organic chemicals that are referred to as *anthropogenic* (`due to humans in nature´) or *xenobiotic* (`foreign to organisms´) chemicals. In this chapter we aim to clarify the key properties and functionalities of the most common groups of pollutants as a result from human activities, and provide some background on how we can group them and understand their behaviour in the environment.

In the field of environmental toxicology, we are most often concerned about the effects of two distinct types of contaminants: metals and organic chemicals. In some cases other chemicals, such as radioactive elements, may also be important while we could also consider the ecological effects of highly elevated nutrient concentrations (eutrophication) as a form of environmental toxicology.

Metals

Metals and metalloids (elements intermediate between metals and non-metals) comprise the majority of the known elements. They are mined from minerals and used in an enormous variety of applications either in their elemental form or as chemicals with inorganic or organic elements or ions. Many metals occur as cations, but many processes influence the dissolved form of metals.

Aluminium for example for example is only present under very acidic conditions as dissolved cation (Al^{3+}), while at neutral pH the metal speciates into for example certain hydroxides ($Al(OH_3)^0$). Mercury as free ion is present at (Hg^{2+}) but due to microbial transformation the highly toxic product methylmercury (CH_3Hg^+, or $MeHg^+$) is formed. Mining and processing of metals together with disposal of metal-containing wastes are the main contributors to metal pollution although sometimes metals are introduced deliberately into the environment as biocides. The widely used pesticide copper sulfate in e.g. grape districts is an example (LINK on comparison to glyphosate here). More information on metals considered to be environmental pollutants is given in Section 2.2.1.

Organic chemicals

Organic chemicals are manufactured to be used in a wide variety of applications. These range from chemicals used as pesticides to industrial intermediates, fossil fuel related hydrocarbons, additives used to treat textiles and polymers, such as flame retardants and plasticisers, and household chemicals such as detergents, pharmaceuticals and cosmetics.

Organic chemicals that we regard as environmental pollutants include a huge variety of different structures and have a wide variety of properties that influence environmental distribution and toxicity. With such a wide variety of chemicals to deal with, it is useful to classify them into groups. Depending on our interest, we can base this classification on different aspects, for example on their chemical structure, their physical and chemical properties, the applications the chemicals are used in, or their effects on biological systems. These aspects are of course closely related to their chemical structure as this is the basis of the properties and effects of chemicals. An overview of different ways of classifying environmental contaminant (sometimes referred to as ecotoxicants) is shown in Tables 1A, 1B, and 1C.

Table 1A. *Grouping options of organic contaminants with specific chemical structures*

Term	Characteristics	Examples
Hydrocarbons	More CH_x units: higher hydrophobicity/lipophilicity, and lower aqueous solubility	hexane
Polycyclic aromatic hydrocarbons	Combustion products. Flat structure	naphthalene, B[a]P
Halogenated hydrocarbons	H substituted by fluor, chlorine, bromide, iodine. Often relatively persistent	PCB, DDT, PBDE
Dioxins and furans	Combustion/industrial products, one or two oxygen atoms between two aromatic rings. Highly toxic.	TCDD, TCDF
Organometallics	Organic chemicals containing metals, used e.g. in anti-fouling paints	tributyltin
Organophosphate pesticides	Phosphate esters, often connecting two lipophilic groups. Act on nervous system	chlorpyrifos
Pyrethroids	Usually synthetic pesticides based on natural pyrethrum extracts	fenvalerate
Neonicotinoids	Synthetic insecticides with aromatic nitrogen, related to the alkaloid nicotine	imidacloprid
… Endless varieties / combinations and ...	...too many characteristics to list	...

Table 1B. *Grouping options of organic contaminants with specific properties*

Term	Characteristics	Examples
Persistent organic pollutants (POPs)	Bioaccumulative, end up even in remote Arctic systems	PCBs, PFOS
Persistent mobile organic chemicals (PMOCs)	Difficult to remove during drinking water production	PFBA, metformin
Ionogenic organic chemicals (IOCs)	Acids or bases, predominantly ionized under environmental pH	Prozac, MDMA, LAS
Substances of unknown or variable composition, complex reaction products or of biological materials (UVCB)	Multicomponent compositions of often analogue structures with wide ranging properties.	Oil based lubricants
Plastics	Chains of repetitive monomer structures. Wide ranging size/dimensions.	Polyethylene, silicone, teflon
Nanoparticles (NP)	Mostly manufactured particles with >50% having dimensions ranging 1 - 100 nm.	Titanium dioxide (TiO_2), fullerene

Table 1C. *Grouping options of organic contaminants with specific usage*

Term	Characteristics	Examples
Pesticides **Herbicides** **Insecticides** **Fungicides** **Rodenticides** **Biocides**	Toxic to pests Toxic to plants Toxic to insects Toxic to fungi Toxic to rodents Toxic to many species	DDT atrazine, glyphosate Chlorpyrifos, parathion Phenyl mercury acetate Hydrogen cyanide Benzalkonium
Pharmaceuticals	Specifically bioactive chemicals with often (un)known side effects. Many bases.	diclofenac (pain killer), iodixanol (radio-contrast), carbemazepine, prozac
Drugs of abuse	Often opioid based but also synthetic designer drugs with similar activity. Many are are ionogenic bases.	cannabinoids, opioids, amphetamine, LSD
Veterinary Pharmaceuticals	Can include relatively complex (ionogenic) structures	antibiotics, antifungals, steroids, non-steroidal anti-inflammatories
Industrial Chemicals	Produced in large volumes by chemical industry for a wide array of products and processes	phenol
Fuel products	Flammable chemicals	kerosene
Refrigerants and propellants	Small chemicals with specific boiling points	freon-22
Cosmetics/personal care products	Wide varieties of specific ingredients of formulations that render specific properties of a product	sunscreen, parabenes
Detergents and surfactants	Long hydrophobic hydrocarbon tails and polar/ionic headgroups	Sodium lauryl sulfate (SLS), benzalkonium
Food and Feed Additives	To preserve flavor or enhance its taste, appearance, or other qualities	"E-numbers", acetic acid = E260 in EU, additive 260 in other countries

Chapter 2 mostly discusses groups of chemicals in separate modules according to the *specific environmental properties in Table 1B* (Section 2.2) and *specific applications in Table 1C* (Section 2.3), according to which certain regulations apply in most cases. The property classifications can be based on (often interrelated) properties such as solubility (in water), hydrophobicity (tendency to leave the water), surface activity (tendency to accumulate at surfaces of two phases, such as for "surfactants"), polarity, neutral or ionic chemicals and reactivity. Other classifications very important for environmental toxicology are based on environmental behaviour or effects, such as persistency ("P" increasing problems with increased emissions), bioaccumulation potential ("B", up-concentration in food chains), or type of specific toxic effects ("T"). The influence of ***specific chemical structures*** such as in Table 1A is further clarified in the current introductory chapter in order to better understand the basic chemical terminology.

Structures of organic chemicals and functional groups

- *Hydrocarbons and polycyclic aromatic hydrocarbons*

As the name suggest, hydrocarbons contain only carbon and hydrogen atoms and can therefore be considered to be the simplest group of organic molecules. Nevertheless, this group covers a wide variety of aliphatic, cycloaliphatic and aromatic structures (see Figure 1 for some examples) and also a wide range of properties. What this group shares is a low solubility in water with the larger molecules being extremely insoluble and accumulating strongly in organic media such as soil organic matter.

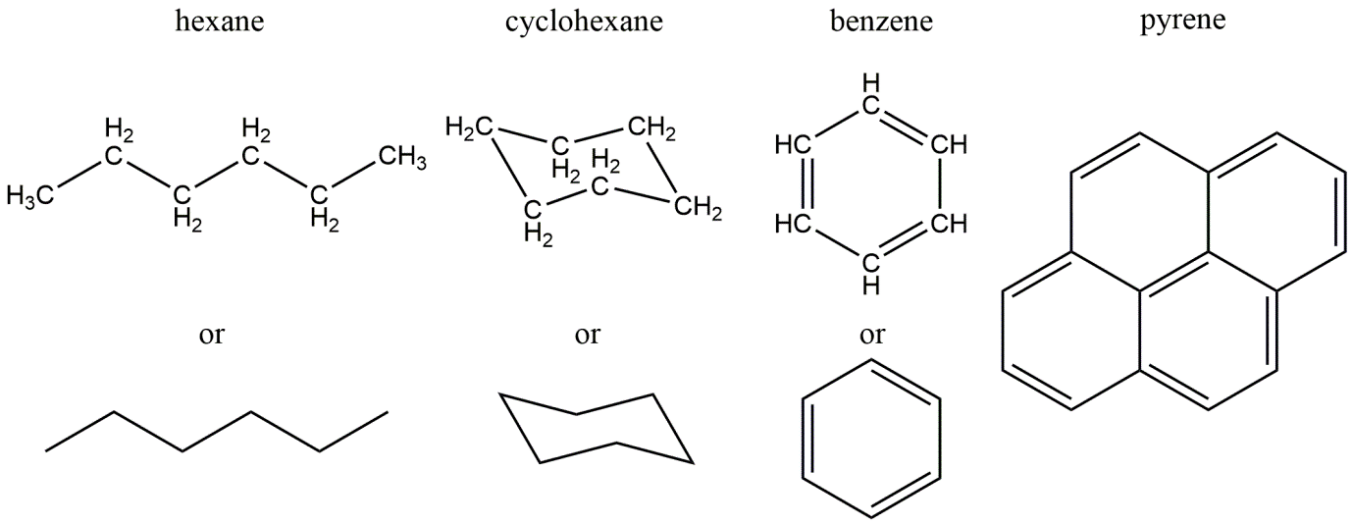


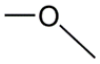
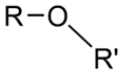
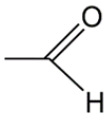
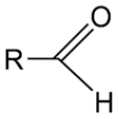
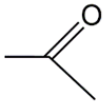
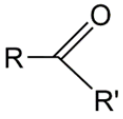
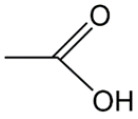
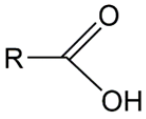
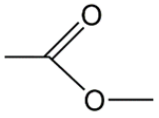
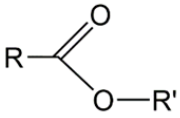
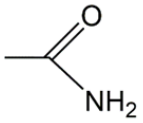
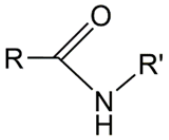
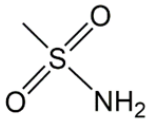
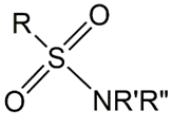
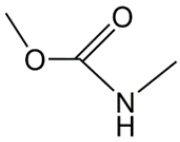
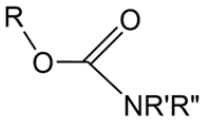
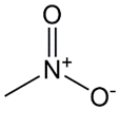
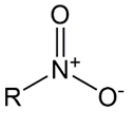
Figure 1. Examples of hydrocarbons.

As a result of the ability of carbon to form strong bonds with itself and other atoms to form structures containing long chains or rings of carbon atoms there is a huge and increasing number (millions) of organic chemicals known. Chemicals containing only carbon and hydrogen are known as hydrocarbons. Aliphatic molecules consist of chains of carbon atoms as either straight or branched chains. Molecules containing multiple carbon-carbon bonds (C=C) are known as unsaturated molecules and can be converted to saturated molecules by addition of hydrogen.

Cyclic alkanes consist of rings or carbon atoms. These may also be unsaturated and a special class of these is known as aromatic hydrocarbons, for example benzene in Figure 1. The specific electronic structure in aromatic molecules such as benzene makes them much more stable than other hydrocarbons. Multiple aromatic rings linked together make perfectly flat molecules, such as pyrene in Figure 1, that can be polarized to some extent because of the shared electron rings. In larger sheets, these polycyclic aromatic molecules also make up the basic graphite structure in pencils, and also typically represent the strongly adsorbing surfaces of black carbon phases such as soot and activated carbon.

The structures of organic chemicals help to determine their properties as behaviour in the environment. At least as important in this regard, however, is the presence of functional groups. These are additional atoms or chemical groups that are present in the molecule that have characteristic chemical effects such as increasing or decreasing solubility in water, giving the chemical acidic or basic properties or other forms of chemical reactivity. The common functional groups are shown in Table 2.

Table 2. *Common Functional Groups, where R are carbon backbone or hydrogen units*

Class	Functional group	General Formula	Examples
Halohydrocarbons	—X (F, Cl, Br, I)	R—X	CH_3I (iodomethane)
Alcohols	—OH	R—OH	CH_3OH (methanol)
Ethers			CH_3OCH_3 (dimethylether)
Aldehydes			CH_2O (methanal)
Ketones			CH_3COCH_3 (propanone, or acetone)
Carboxylic Acids			CH_3COOH (ethanoic acid, or acetic acid)
Esters			$CH_3COOCH_2CH_3$ (ethyl acetate)
Amines	—NH_2	R—NR'R"	CH_3NH_2 (methylamine, or aminomethane)
Amide			$CH_3CON(CH_3)_2$ (N,N-dimethylacetamide)
Sulphonamide			CH_3SOONH_2 (methylsulfonamide)
Carbamate			X
Nitro			CH_3NOO (nitromethane)

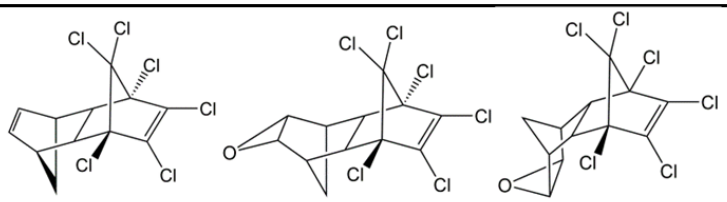
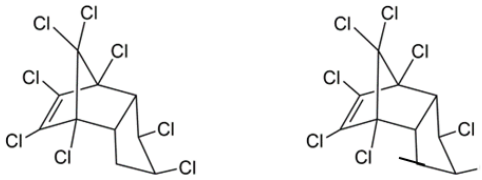
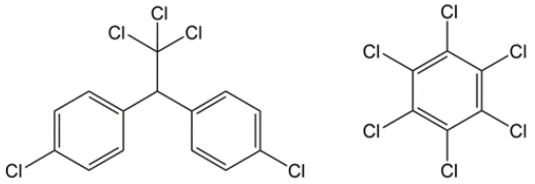
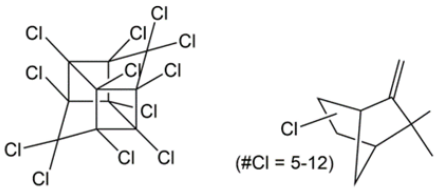
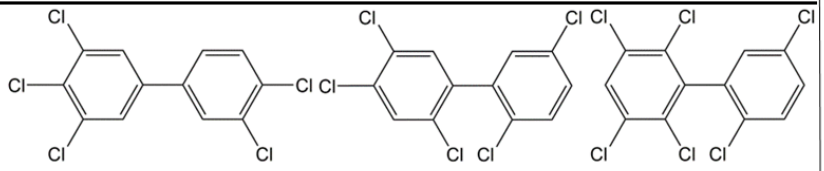
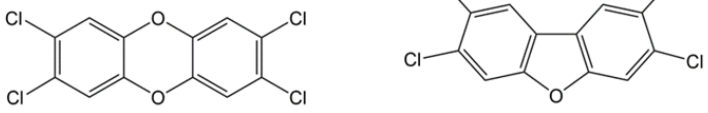
- *Halogenated hydrocarbons: first generation pesticides*

The first organic chemical recognised as an environmental pollutant was the insecticide DDT (see clip 1 at the start of this chapter). It later became clear that other organochlorine pesticides such as lindane and dieldrin (Table 3) were also widely distributed in the environment. This was also the case for polychlorinated biphenyls (PCBs) and other organochlorinated industrial chemicals. These chemicals all share a number of undesirable properties such as environmental persistence, very low solubility in water and high level of accumulation in biota to potentially toxic levels. Many organochlorines can be viewed as hydrocarbons in which hydrogen atoms have been replaced by chlorine. This makes them even less soluble than the corresponding hydrocarbon due to the large size of chlorine atoms. In addition, chlorination also makes the molecules more chemically stable and therefore contributes to their environmental persistence. Other organochlorines contain additional functional groups, such as the ether bridges in PCDDs and PCDFs (better known as dioxins and dibenzofurans) and ester groups in the 2,4-D and 2,4,5-T herbicides. Many organochlorines were applied very successfully in huge quantities as pesticides for decades before their negative effects such as persistence and accumulation in biota became apparent. It is therefore no coincidence that the initial set of Persistent Organic Pollutants (POPs) identified in the Stockholm Treaty (see below) as chemicals that should be banned were all organochlorines, as shown in Table 3.

As well as chlorine, other halogens such as bromine and fluorine are used in important groups of environmental contaminants. Organobromines are best known as flame retardants and have been applied in large quantities to improve the fire safety of plastics and textiles. They share many of the same undesirable properties of organochlorines and several classes have now been taken out of production. Organofluorines are another important class of halogenated chemicals, and part of the well-known group of ozone depleting CFCs (Section 2.3.6). In particular, per-and polyfluoralkyl substances are widely used as fire-stable surfactants in fire-fighting foams, as grease and water resistant coatings and in the production of fluoropolymers such as Teflon. Organofluorines are much more water soluble and much less bioaccumulative than organochlorines and organobromines but are extremely persistent in the environment.

The recognition of these organochlorines as harmful environmental contaminants eventually resulted in measures to restrict their manufacture and use in the Stockholm Convention on Persistent Organic Pollutants signed in 2001 to eliminate or restrict the production and use of persistent organic pollutants (POPs). This initial list of POPs has been subsequently augmented with other harmful halogenated organic pollutants up to a total of 29 chemicals, which are either to be eliminated, restricted, or required measured to reduce unintentional releases. POPs are further discussed in section 2.2.4.

Table 3. *Key persistent organic pollutants, also named POPs - the Dirty Dozen*

Chemical	Use	Structure
Aldrin/ Dieldrin/ Endrin	crop insecticides (corn, cotton, grains)	
Chlordane/ Heptachlor	insecticides (potatoes, citrus vegetables, cotton/termites, soil insects)	
DDT/ Hexachloro -benzene	crop insecticide (cotton) / fungicide for seed treatment	
Mirex/ Toxaphene	insecticide (termites, fire ants/ livestock and crops)	
PCBs	industrial (heat exchange fluid for e.g. electrical transformers, paint and plastic additive)	 PCB 81 (co-planar structure) PCB 101 PCB 151 (non-planar structures)
Dioxins/ Furans	unintentionally produced during combustion	 TCDD (2,3,7,8 tetrachlorodioxin) TCDF (2,3,7,8 tetrachlorofuran)

Additional POPs to eliminate include: chlordecone, lindane (hexachlorocyclohexane), pentachlorobenzene, endosulfan, chlorinated naphthalenes, hexachlorobutadiene, tetrabromodiphenylether, and pentabromodiphenylether decabromodiphenyl ether (BDEs).

- *Alternatives for the organochlorine pesticides: effective functional groups*

Since the signing of the Stockholm Convention, organochlorine pesticides have been replaced in most countries by more modern pesticide types such as the organophosphorus and carbamate insecticides. These compounds are less persistent in the environment, but still could pose elevated risks to environments surrounding the agricultural sites, and increased levels on food produced on these agricultural sites. The very toxic organophosphorus neurotoxicant parathion has been in use since the 1940s, and has the typical two lipophilic side chains on two esters (ethyl units), as well as a polar unit. Parathion has caused hundreds of fatal and non-fatal intoxications worldwide and as a result it is banned or restricted in 23 countries. The relatively comparable organophosphate structure of diazinon has been widely used for general-purpose gardening and indoor pest control since the 1970s, but residential use was banned in the U.S. in 2004. In Californian agriculture however, 35000 kg diazinon was used in 2012. The carbamate based insecticide carbaryl is toxic to target insects, and also non-target insects such as bees, but is detoxified and eliminated rapidly in vertebrates, and not secreted in milk. Although illegal in 7 countries, carbaryl is the third-most-used insecticide in the U.S., approved for more than 100 crops. In 2012, 52000 kg carbaryl was used in California, while this was 3 times more in 2000. Neonicotinoid insecticides, with the typical aromatic ring containing nitrogen, form a third generation of pesticide structures. Imidacloprid is currently the most widely used insecticide worldwide, but as of 2018 banned in the EU, along with two other neonicotinoids clothianidin and thiamethoxam.

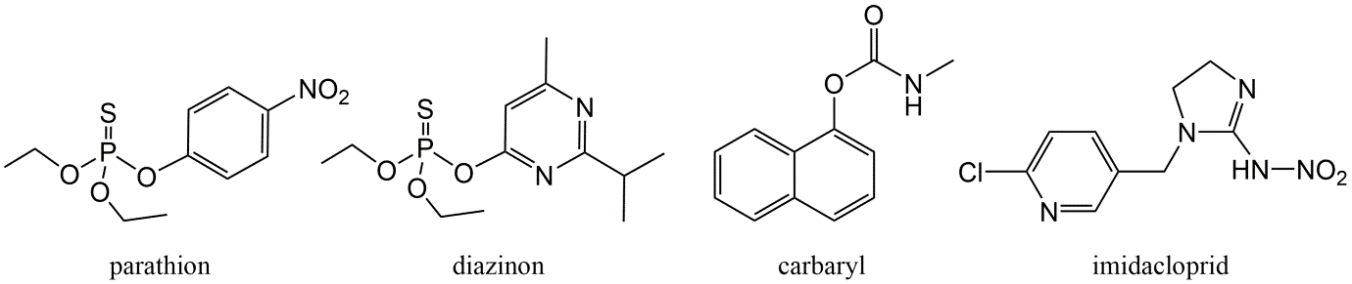


Figure 2. Some examples of second and third generation replacements of organochlorine pesticides.

- *Relatively simple and (very) complex pollutants*

As well as the pesticides discussed above, many other chemicals are brought into the environment inadvertently during their manufacture, distribution and use and the range of chemicals recognised as problematic environmental contaminants has expanded enormously. These include fossil fuel-related hydrocarbons, surfactants, pigments, biocides and chemicals used as pharmaceuticals and personal care products (PPCPs). Figure 1 in Boxall et al. (2012) gives an illustrative overview of the major routes by which PPCPs, but also many other anthropogenic contaminants other than pesticides, are released into the environment. Particularly wastewater treatment systems form the main entry point for many industrial and household products.

The wide variety of contaminant structures does not mean that most chemicals have become increasingly more complex. For risk assessment, molecular properties such as water solubility, volatility and lipophilicity are often estimated based on quantitative structure-property relationships (Section 3.4.3). With increasingly complex structures, such property-estimations based on the molecular structure become more uncertain.

The antibiotic erythromycin for example (Figure 3), is a very complex chemical structure ($C_{37}H_{67}NO_{13}$) that has 13 functional units along with a 14 member ring. In addition, the tertiary nitrogen group is an amine base group that can give the molecule a positive charge upon protonation, depending on the environmental pH. Erythromycin is on the World Health Organization's List of Essential Medicines (the most effective and safe medicines needed in a health system), and therefore widely used. Continuous emissions in waste streams pose a potential threat to many ecosystems, but many environmentally and toxicologically relevant properties are scarcely studied, and poorly estimated.

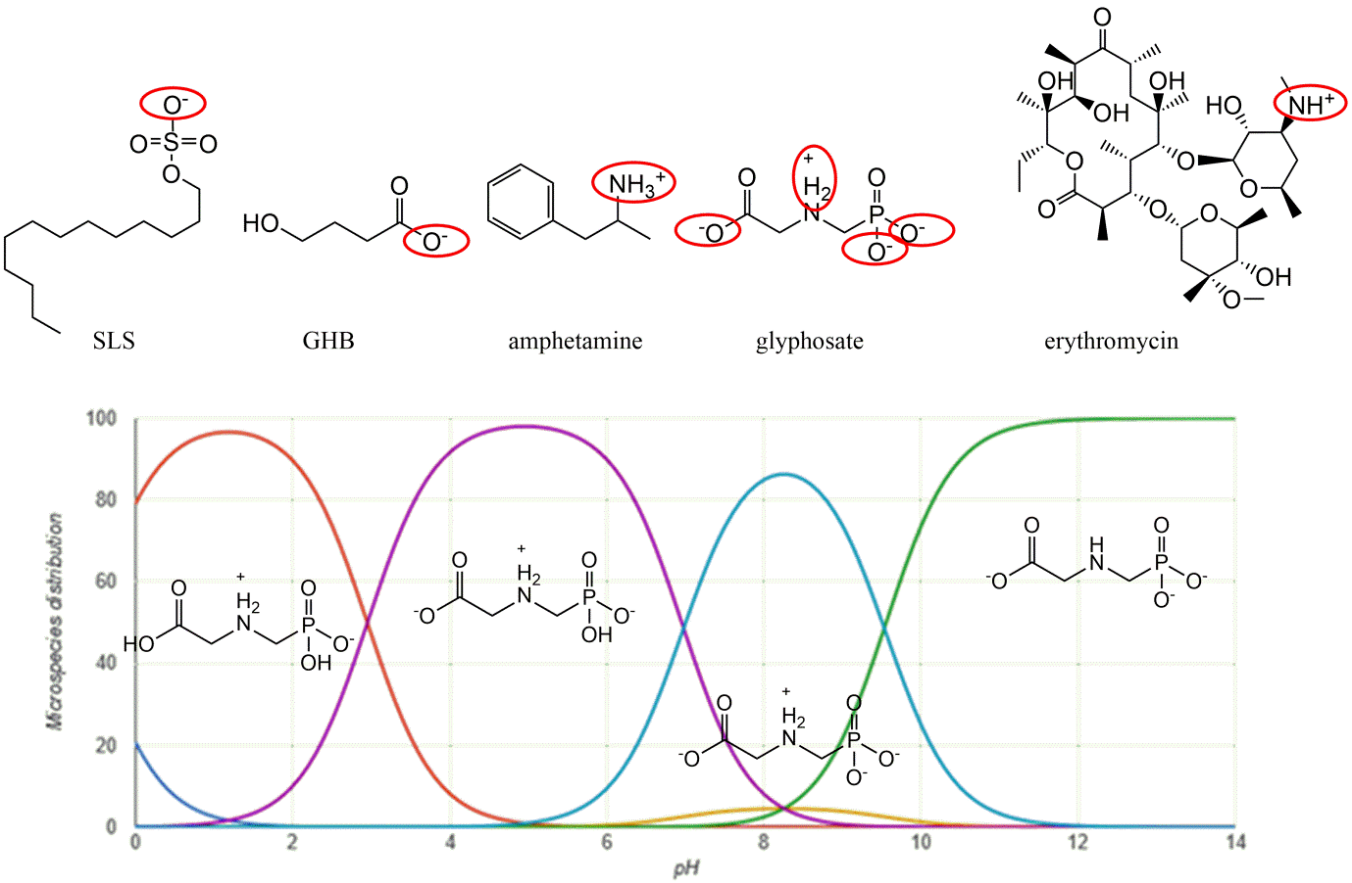


Figure 3. Relatively simple or (very) complex chemicals? Glyphosate speciation profile (chemicalize.org) for the 4 dominant glyphosate different species. Glyphosate in soil (pH4-8) predominantly occurs with 3 charged groups (net charge -1), partly with 4 charged groups (-2). The soap SLS is always negatively charged, GHB predominantly negative (pK_a 4.7), amphetamine predominantly positive (pK_a 9.9), erythromycin predominantly positive (pK_a 8.9)

There are also many contaminants or toxicants with a seemingly simple structure. Many surfactants are simple linear long chain hydrocarbons with a polar or charged headgroup (Figure 4). The illicit drug amphetamine has only a benzene ring and an amine unit, the illicit drug GHB only an alcohol and a carboxylic acid, the herbicide glyphosate only 16 atoms. Still, these 4 chemical examples also have acidic or basic units that often result in predominantly charged organic molecules, which also strongly influences their environmental and toxicological behaviour (see sections on PMOCs and Ionogenic Organic Compounds). In case of glyphosate, the chemical has 4 differently charged forms depending on the pH of the environment. At common pH of 7-9, glyphosate has all charged groups predominantly ionized, making it very difficult to derive calculations on environmental properties.

References

Boxall, A. B. A., Rudd, M. A., Brooks, B. W., Caldwell, D. J., Choi, K., Hickmann, S., Innes, E., Ostapyk, K., Staveley, J. P., Verslycke, T., Ankley, G. T., Beazley, K. F., Belanger, S. E., Berninger, J. P., Carriquiriborde, P., Coors, A., DeLeo, P. C., Dyer, S. D., Ericson, J. F., Gagné, F., Giesy, J. P., Gouin, T., Hallstrom, L., Karlsson, M. V., Larsson, D. G. J., Lazorchak, J. M., Mastrocco, F., McLaughlin, A., McMaster, M. E., Meyerhoff, R. D., Moore, R., Parrott, J. L., Snape, J. R., Murray-Smith, R., Servos, M. R., Sibley, P. K., Oliver Straub, J., Szabo, N. D., Topp, E., Tetreault, G. R., Trudeau, V. L., Van der Kraak, G. (2012). Pharmaceuticals and personal care products in the environment: what are the big questions? Environmental Health Perspectives 120, 1221-1229.

This page titled 2.3.2.1: Introduction is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **2.1: Introduction** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.2.2: Pollutants with specific properties

2.2. Pollutants with specific properties

2.2.1. Metals and metalloids

Author: Kees van Gestel

Reviewers: John Parsons, Jose Alvarez Rogel

Learning objectives:

You should be able to:

- describe the difference between metals and metalloids
- describe a classification using different binding affinities of metals to macromolecules and infer on its importance for their toxicity and/or bioaccumulation
- mention important sources of metal pollution

Keywords: Heavy metals, Metalloids, Rare earth elements, Essential elements

Introduction

The majority of the elements in the periodic table consists of metals: Figure 1.

1	2	3		4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Hydrogen 1 H																		Helium 2 He
Lithium 3 Li	Beryllium 4 Be												Boron 5 B	Carbon 6 C	Nitrogen 7 N	Oxygen 8 O	Fluorine 9 F	Neon 10 Ne
Sodium 11 Na	Magnesium 12 Mg												Aluminum 13 Al	Silicon 14 Si	Phosporus 15 P	Sulfur 16 S	Chlorine 17 Cl	Argon 18 Ar
Potassium 19 K	Calcium 20 Ca	Scandium 21 Sc		Titanium 22 Ti	Vanadium 23 V	Chromium 24 Cr	Manganese 25 Mg	Iron 26 Fe	Cobalt 27 Co	Nickel 28 Ni	Copper 29 Cu	Zinc 30 Zn	Gallium 31 Ga	Germanum 32 Ge	Arsenic 33 As	Selenium 34 As	Bromine 35 Br	Krypton 36 Kr
	Strontium 38 Sr	Yttrium 39 Y		Zirconium 40 Zr			Molybdenum 42 Mo				Silver 47 Ag	Cadmium 48 Cd		Tin 50 Sn	Antimony 51 Sb	Tellurium 52 Te	Iodine 53 I	Xenon 54 Xe
Cesium 55 Cs	Barium 56 Ba	Lanthanum 57 Ln	*							Platinum 78 Pt	Gold 79 Au	Mercury 80 Hg		Lead 82 Pb		Pollonium 84 Po		Radon 86 Rn
	Radium 88 Ra	Actinium 89 Ac	**															

REE * Lanthanides 58-71 ** Actinides 90-103 incl. Uranium 92 U Plutonium 94 Pu

Figure 1. Periodic table of elements, with the most important elements for Environmental Toxicology shown. The shaded elements are metals, the partially shaded elements are metalloids. Bold lettered metals area heavy metals (specific density > 5 g/cm^3). Elements shown within bold lines (and in italics) are essential elements. The Lanthanides and Actinides together are the rare earth elements (REEs). (Source: Steven Droge).

The distinction between metals and heavy metals (density relative to water < or >5 g cm^{-3}) is not very meaningful for such a heterogeneous group of elements with rather different biological and chemical properties. The rare earth elements (REEs), lanthanides and actinides, have, for example, a high density or specific weight but are usually not considered heavy metals because of their rather different chemical behaviour. Metalloids have both metallic and non-metallic properties or are nonmetallic elements that can combine with a metal to produce an alloy. Figure 1 shows the periodic table of elements, indicating the groups of (heavy) metals, metalloids and rare earth elements.

Also indicated in Figure 1 are the elements that are known to be essential to life and include besides C, H, O and N, the major essential elements Ca, P, K, Mg, Na, Cl and S, the trace elements Fe, I, Cu, Mn, Zn, Co, Mo, Se, Cr, Ni, V, Si, As and B (the latter only for plants) and some elements that may support physiological functions at ultra-trace levels (Li, Al, F and Sn) (Walker et al., 2012).

Chemical and physical properties

Except for mercury, most pure metals are solid at room temperature. In general, metals are good electrical and thermal conductors having high luster and malleability. Upon heating, metals readily emit electrons. These descriptors of metals, however, are not very helpful when having to deal with elements that do not exist prominently in the pure elemental state, but rather are present as metal compounds, complexes, and ions at fairly low environmental concentrations.

More useful are characteristics that influence metal transport between environmental compartments and their interaction with abiotic and biotic components of the environment. The speciation, the chemical form in which an element occurs (e.g., oxidized, free ion or complexed to inorganic or organic molecules), determines its transport and interaction in the environment (see Section on Metal Speciation). Chemical bonding is determined by outer orbital electron behavior, with metals tending to lose electrons when reacting with nonmetals. In many normal biological reactions, metals are cofactors within coenzymes (e.g. in vitamins) and can act as electron acceptors and donors during oxidation and reduction reactions (Newman, 2015).

Nieboer and Richardson (1980) proposed a classification, based on the equilibrium constant for the formation of metal complexes. They distinguished:

- **Class A-metals**: acting as hard Lewis acids (electron acceptors) with high affinity for oxygen-containing groups in macromolecules, such as carboxyl and alcohol groups. Al, Ba, Be, Ca, K, Li, Mg, Na and Sr belong to this group;
- **Class B-metals**: acting as soft Lewis acids with high affinity for nitrogen- and sulphur-containing groups in macromolecules, such as amino and sulphydryl groups. This group includes Ag, Au, Bi, Hg, Pd, Pt and Tl.

In addition, an intermediate or **borderline** group is defined, in which the type A or B characteristics are less pronounced. As, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, Sb, Sn, Ti, V, and Zn belong to this group.

This classification of metals is highly relevant for the transport across cell membranes, the intercellular storage in granules and the induction of metal-binding proteins as well as for their behaviour in the environment in general.

Occurrence

(Heavy) metals and rare earth elements are diffusely distributed over the Earth, but at some places certain elemental combinations are highly concentrated (in metal ores). Despite this diffuse distribution, differences in background metal concentrations in soils can be large, depending on the type and origin of rock or sediment (Table 1).

Table 1. *Background concentrations (mg/kg dry weight) of (heavy) metals and metalloids in crust material and median and maximum concentrations in different top soils across the world. Derived from Kabata-Pendias and Mukherjee (2007) and Alloway (2013).*

Element	Upper crust	Median	Maximum
Ag	0.06-0.07	0.05-1.0	3.15
As	1.8-2.0	1.9-7.3	123
Be	3	0.92-1.9	
Cd	0.1	0.13-0.7	41
Co	10-12	2.4-18	347
Cr	35-100	9.9-96	5,910
Cu	14-55	7-109	3,088
Hg	0.07	0.037-0.086	3.5
La	30	23-37	
Li	20	13-28	
Mn	527-900	255-752	42,603
Mo	1.4-1.5	<0.5-3.0	74
Ni	19-20	4.9-50	4,955
Pb	17	8.4-24.8	2,635
Se	0.05	0.14-0.74	17
Sn	2.5	<1.0-1.68	110
Ti	4,400	2,900-15,480	
Tl	0.5	0.23-0.66	24
V	53-135	18-320	490
Zn	52-70	26.8-89	6,173
Zr	165	92-421	

In general, volcanic rock (e.g. basalt) contains high and sedimented rock (e.g. limestone) low metal levels. But there is no relation between metal concentrations in the Earth's crust and the elemental requirements of organisms.

Emissions of metals

Upon weathering of stone formations and ores, elements are released and enter local, regional and global biogeochemical cycles. Depending on their water solubility and on soil properties and vegetation, metals may be transported through the environment and deposited or precipitated at places close to or far away from their source.

Volcanoes take account of the largest natural input of metals to the environment but the concentrations of these metals in the soil are rarely elevated to toxic levels due to the massive dilution which takes place in the atmosphere. Permanently active volcanoes may be an important local source of (metal) pollution.

A special case is arsenic, that may occur as a natural element of soils. At some places, As levels are fairly high, particularly in ground water. High As-groundwater areas are found in Argentina, Chile, Mexico, China and Hungary, and also in Bangladesh, India (West Bengal), Cambodia, Laos and Vietnam. In the latter countries, especially in the Bengal Basin, millions of wells have been dug to provide safe drinking water. Irrigation pumping leads to an inflow of oxygen and organic carbon, which causes a mobilisation of arsenic normally bound to ferric oxyhydroxides in these soils. As a result in many wells dissolved As concentrations are exceeding the World Health Organisation (WHO) guideline value of 10 µg/L for drinking water.

Important anthropogenic sources of metals in the environment include:

- *Metal mining*, which may also lead to an enormous physical disturbance of the environment (destruction of ecosystems).
- *Metal smelting*.
- Use of metals in *domestic and industrial products*, and the discharge of domestic waste and sewage.
- *Metal-containing pesticides*, e.g. 'Bordeaux Mixture (copper sulphate with lime ($Ca(OH)_2$), used as a fungicide in viniculture, hop-growing and fruit-culture, and metal-containing fungicides, such as organo-tin compounds.
- The use of metals and especially of REEs in *microelectronics*.
- *Energy-producing industries* burning coal and oil, and producing metal-containing fly ash.
- *Transport of energy and traffic* making use of electricity, giving rise to corrosion of electric wires and pylons.
- *Non-metal industries*, e.g. leather (chromium) and cement production (thallium).
- , using Tetra Ethyl Lead (TEL) as anti-knocking agent in petrol (nowadays banned in most countries) and the use of catalysts in cars (platinum, palladium).

Anthropogenic releases of many metals, such as Pb, Zn, Cd and Cu, are estimated to be between one and three orders of magnitude higher than natural fluxes (Depledge *et al*. 1998). An estimated amount of up to 50,000 tonnes of mercury are released naturally per year as a result of degassing from the Earth's crust, but human activities account for even larger emissions (Walker *et al*. 2012).

References

Alloway, B.J. (2013). Heavy Metals in Soils. Trace Metals and Metalloids in Soils and their Bioavailability. Third Edition. Environmental Pollution, Volume 22, Springer, Dordrecht.

Depledge, M.H., Weeks, J.M., Bjerregaard, P. (1998). Heavy metals. In: Calow, P. (Ed.). Handbook of Ecotoxicology. Blackwell Science, Oxford, pp. 543-569.

Kabata-Pendias, A., Mukherjee, A.B. (2007). Trace Elements from Soil to Human. Springer Verlag, Berlin.

Newman, M.C. (2015). Fundamentals of Ecotoxicology. The Science of Pollution. Fourth Edition. CRC Press, Taylor & Francis Group. Boca Raton.

Nieboer, E., Richardson, D.H.S. (1990). The replacement of the nodescript term 'Heavy metals' by a biologically and chemically significant classification of metal ions. Environmental Pollution (Ser. B) 1, 3-26.

Walker, C.H., Hopkin, S.P., Sibly, R.M., Peakall, D.B. (2012). Principles of Ecotoxicology, Fourth Edition. CRC Press Taylor & Francis Group, London.

2.2. Question 1

Why can the term heavy metal not be used when referring to different elements considered metals?

2.2. Question 2

Why are some elements indicated as essential elements?

2.3. Question 3

Why is the classification of Nieboer and Richardson relevant for the biological interactions of metals with living organisms?

2.4. Question 4

Name at least 5 sources of metal emission to the environment.

2.5. Question 5

Besides leading to the emission of metals, metal mining may another important environmental effect. Which one?

2.2.2. Radioactive compounds

in preparation

2.2.3. Industrial Chemicals

Authors: Steven Droge

Reviewer: Michael McLachlan

Leaning objectives:

You should be able to

- discuss a history perspective on key chemical legislations around the world
- look up registration dossiers yourself to obtain relevant ecotoxicological information
- realize that complete dossiers are most urgent for high production tonnage substances and the most hazardous substances
- understand why for some groups of chemicals already specific regulations were in place apart from common industrial substances.

Keywords: Chemical industry, tonnage, hazardous chemicals, REACH, regulation

Introduction

The chemical industry produces a wide variety of chemicals that find use in industrial process and as ingredients in day-to-day products for consumers. Instead of chemicals, 'substances' may be a more carefully worded description as it also includes complex mixtures, polymers and nanoparticles. Many substances are produced by globally distributed companies in very high volumes, ranging for example from 100 - 10,000 tonnes (1 tonne = 1000 kg) per year. Worldwide, governments have tried to control and assess chemical safety, as nicely summarized on the ChemHAT website. Australia for example, has the *Industrial Chemicals (Notification and Assessment) Act 1989* (2013 version). Just like elsewhere in the world, in the European Union (EU) a variety of regulatory institutes at all levels of government used to perform safety assessments regarding the use of substances in products, and how these are emitted into waste streams. This changed dramatically in 2007.

On June 1st, 2007 (Figure 1), a new EU regulation went into force called REACH (official legislation documents C 1907/2006; about REACH; EU info on REACH). This law reversed the role of governments in chemical safety assessment, because it placed the burden of proof on companies that manufacture a chemical, import a chemical into the EU, or apply chemicals in their products. Within REACH companies must identify and manage the risks linked to the chemicals they manufacture and market in the EU. REACH stands for Registration, Evaluation, Authorisation and Restriction of Chemicals. China soon followed with the analogous "China REACH" in 2010, and then came South Korea in 2015 with "K-REACH". The main focus in this module is on EU-REACH as the leading and well documented example. Other legislation regulating industrial chemicals can often be easily found online, e.g. via the ChemHAT link above.

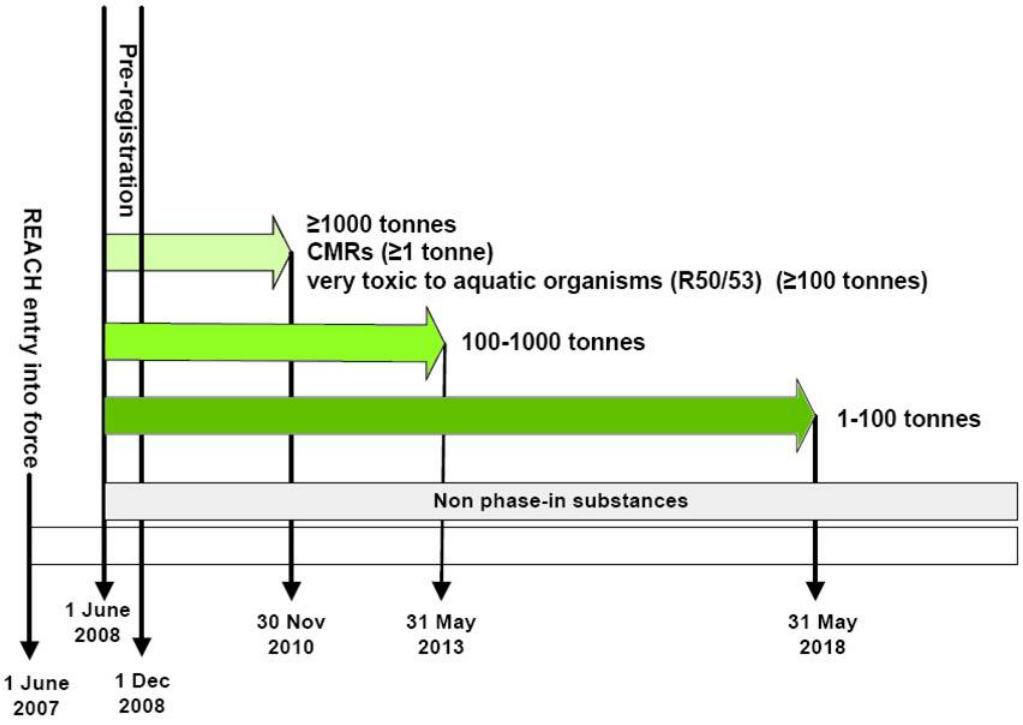


Figure 1. Scheme for the registration phase of the REACH regulation for existing industrial chemicals of different tonnage bands and hazardous potential, as well as newly developed chemicals ("Non phase-in") for the EU market. CMRs = chemicals that are proven carcinogenic, mutagenic or toxic to reproduction. R50/R53 labels indicate "Very toxic to aquatic organisms"/ "May cause long-term adverse effects in the aquatic environment". Source: http://www.cirs-reach.com/REACH/REACH_Registration_Deadlines.html (with permission).

In REACH, each chemical is registered only once. Accordingly, companies must work together to prepare one dossier that demonstrates to the European Chemical Agency (ECHA) how chemicals can be safely used, and they must communicate the risk management measures to the users. ECHA, or any Member State, authorizes the dossiers, and can start a "restriction procedure" when they are concerned that a certain substance poses an unacceptable risk to human health or the environment. If the risks cannot be managed, authorities can restrict the use of substances in different ways. In the long run, the most hazardous substances should be substituted with less dangerous ones.

So which chemicals have been registered in the past decade (2008-2018) in REACH?

In principle, REACH applies to all chemical 'substances' in the EU zone. This includes metals, such as "iron" and "chromium", organic chemicals such as "methanol" and "fatty acids" and "ethyl-4-(8-chloro-5,6-dihydro-11H-benzo[5,6]cyclohepta[1,2-

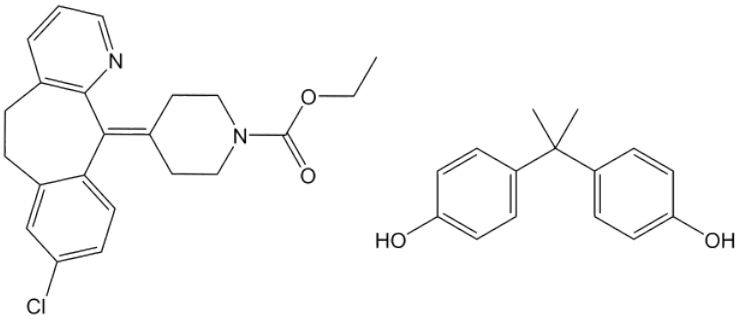
b]pyridin-11-ylidene)piperidine-1-carboxylate (see **Box 1**)", and (nano)particles like "zink oxide" and "silicon dioxide", and polymers. Discover for example the registration dossier link in Box 1.

Box 1. Examples from the REACH dossiers

The REACH registration data base can be searched via **LINK**. Accept the disclaimer, and you are ready to search for chemicals based on name, CAS number, substance data, or use and exposure data.

Search for example for the name "ethyl 4-(8-chloro-5,6-dihydro-11H-benzo[5,6]cyclohepta[1,2-b]pyridin-11-ylidene)piperidine-1-carboxylate" and you find the link to the dossier of this substance with CAS 79794-75-5 as compiled by the registrant. This complex chemical name is better known as the antihistamine drug Loratadine, but this name does not show up in the dossier search!

Click on the name to get basic information on the compound. The hazard classification reads: "*Warning! According to the classification provided by companies to ECHA in **REACH registrations** this substance is very toxic to aquatic life, is very toxic to aquatic life with long lasting effects, is suspected of causing cancer, causes serious eye irritation, is suspected of causing genetic defects, causes skin irritation, may cause an allergic skin reaction and may cause respiratory irritation.*" This compound is "PBT" labeled based on limited available data (classifying as a combination of Persistent / Bioaccumulative / Toxic). However, the section [About this substance] reads: "for industrial use resulting in the manufacture of another substance (use of intermediates)." As an intermediate in a restricted process, many parts of the dossier did not have to be completed for REACH. As a medicinal product Loratadine is strictly regulated elsewhere. Scroll down to the REACH link for the registration dossier (.../21649) to find out more the different entries for this chemical.



If we do a search for [" Bisphenol "], we get a long list of optional chemicals, for example Bisphenol A (CAS 80-05-7) but also for example Bisphenol S if you scroll down further (CAS 80-09-1). If we look at the dossier of the first Bisphenol A entry, with tonnage "100 000 - 1 000 000 tonnes per annum", you can find a long list of REACH information packages besides the dossier, as this chemical is hotly debated. The dossier for Bisphenol A was evaluated in 2013, and also this is available (look for the pdf in the Dossier evaluation status). In this compliance check, the registrant is requested to submit additional rat and mouse toxicity data, along with statements of reasons. There is for example also a link to the [Restriction list (annex XVII)], which leads to a pdf called 66.pdf, which states an adopted restriction for this chemical within the REACH framework and the previous legislation, Directive 76/769/EEC: "Shall not be placed on the market in thermal paper in a concentration equal to or greater than 0,02 % by weight after 2 January 2020".

Find your own chemical of interest to discover more on the transparancy of the chemical information on which risk assessment is based.

However, some groups of chemicals are (partly) exempt from REACH because they are covered by other legislation in the EU:

- Active substances used in plant protection products (Section 2.3.1) and biocidal products (Section 2.3.2) are considered as already having been registered and assessed by institutes separate from ECHA. Biocides such as disinfectants and pest control products are per definition hazardous chemicals, but they are also very useful in many ways. The very strict and elaborate biocide laws aim to verify that the potential risk of harm associated with the intended emission scenarios is in balance with expected benefits.
- Food and feedstuff additives (Section 2.3.9) have different legislation and authorisation laws to demonstrate (following a scientific evaluation) that the additive has no harmful effects on human and animal health or on the environment (developed since 1988, schematic graph, Regulation (EC) No 1331/2008)
- Medicinal products (Sections 2.3.3 and 2.3.4) have different legislation and authorisation laws to guarantee high standards of quality and safety of medicinal products, while promoting the good functioning of the internal market with measures that encourage innovation and competiveness (starting with Directive 65/65 in 1965, an overview since, a pdf of the 2001 EU legislation 2001/83/EC)
- "Waste" is not part of the REACH domain, but a product recovered from waste is not.

A detailed overview of European chemical safety guidelines related to chemicals with different application types is presented in Figure 2 in Van Wezel et al. (2017).

Following pre-registration of the 145,297 chemicals most likely to require regulation, REACH came into force in 2008 in a stepwise process with different deadlines for different groups of chemicals. The first dossiers were to be completed by 2010 for the highest produced volume chemicals (>1000 tonnes/y) and the most hazardous chemicals (CMRs >1 tonne/y, and chemicals with known very high aquatic toxicity >100 tonnes/y). These groups potentially pose the greatest risk because of either their high emissions or their inherent toxicity. In 2013, registration dossiers for chemicals with a lower tonnage (100-1000 tonnes/y) were to be completed. By May 31 2018, all chemicals with a quantity of 1-100 tonnes/y chemicals on the EU market should have been registered. New chemicals will all be subject to the REACH procedures.

In 2018, 21.787 substances had been registered under REACH. A total of 14.262 companies were involved. In comparison, 15.500 substances were registered in 2016 (i.e., 6287 chemicals were added in the two following years). In 2018, 48% of all substance registrations had been done in Germany. For 24% of the registered substances a dossier was already available prior to REACH, 70% are "old chemicals" for which no registration had been done before REACH was initiated, and only 6% are newly developed substances that needed to be registered before manufacture or import could start.

There are multiple benefits of REACH regulation of industrial chemicals. Most data on chemicals entered in the registration process is publically available, creating transparency and improving customer awareness. If registered chemicals are classified as Substance of Very High Concern (SVHC) based on the chemical information in these dossiers and after agreement from research panels, alternatives that passed the same regulation can be suggested instead.

The necessity to add data on potential toxicity for so many chemicals has been combined with a strong focus on, and further development of, animal friendly testing methods. Read-across from related chemicals, weight of evidence approaches, and calculations based on chemical structures (QSAR) allow much experimental testing to be circumvented. In vitro studies are also used, but a 2017 REACH document (REACH alternatives to animal testing 2017, which followed up 2011 and 2014 reports) reports that 5.795 *in vitro* studies were used overall to determine endpoints for REACH, compared to 9,287 *in vivo* studies (ratio of 0.6). Clearly, many new animal tests have been performed under REACH to complete the dossiers on industrial chemicals. Prenatal developmental and repeated dose toxicity testing as well as extended one generation reproductive toxicity studies remain difficult to circumvent without animal use. However, the safe use of industrial chemicals must be ensured and demonstrated.

References:

Van Wezel, A.P., Ter Laak, T.L., Fischer, A., Bäuerlein, P.S., Munthee, J., Posthuma, L. (2017). Mitigation options for chemicals of emerging concern in surface waters; operationalising solutions-focused risk assessment . Environmental Science: Water Research 3, 403-414.

2.2.3. Question 1

Name three important advantages of the set-up of an international chemical legislation such as REACH.

2.2.3. Question 2

Name three important challenges, or even disadvantages, of the set-up of an international chemical legislation such as REACH.

2.2.3. Question 3

The total amount of chemical substances for which CAS numbers have been appointed, was 68,062,538 on June 28, 2019. The CAS REGISTRY is updated daily with thousands of new substances. Provide some reasons why only 21,787 substances have been registered for REACH in 2018.

2.2.4. POPs

(draft)

Authors: Jacob de Boer

Reviewer:

Leaning objectives:

You should be able to

- understand how POPs are defined
- recognize chemical structures of POPs
- gain knowledge on the purpose of the Stockholm Convention on POPs

Keywords: Persistence, bioaccumulation, long range transport, toxicity, analysis

Introduction

Chemicals are generally produced because they have a useful purpose. These purposes can vary widely, such as to protect crops by killing harmful insects or fungi, to protect materials against catching fire, to act as a medicine, to enable a proper packing of food materials, etc. Unfortunately, the properties which make a chemical attractive to use, often have a downside when it comes to environmental behavior and/or human health. A number of synthetic chemicals have properties that make them persistent organic pollutants (POPs). POPs are xenobiotic (foreign to the biosphere) chemicals that are persistent, bioaccumulative and toxic ('PBT') in low doses. In addition, they are transported over long distances. Criteria for these properties, which are used to define a chemical as a POP, were set by the United Nations (UN) Stockholm Convention, which was adopted in 2001 and entered into force in 2004 (Fiedler et al., 2019). These criteria are summarized in Table 1 (http://chm.pops.int). The objective of the Stockholm Convention is defined in article 1: "Mindful of the precautionary approach, to protect human health and the environment from the harmful impacts of persistent organic pollutants". Initially, 12 chemicals (aldrin, chlordane, dieldrin, DDT, endrin, heptachlor, hexachlorobenzene (HCB), mirex, polychlorinated biphenyls (PCBs), polychlorinated dibenzo-*p*-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs) and toxaphene) were listed as POPs. Gradually the list was extended with new POPs, which appeared to fulfil the criteria. For some of the new chemicals exceptions were made for limited use, in case no suitable alternatives are available. For example, in the battle against malaria DDT can still be used to a limited extent for in-house spraying in Africa (Van den Berg, 2009). Until now all POPs are chemicals that contain carbon and halogen atoms. Some POPs, such as the PCDDs and PCDFs (together often short-named as dioxins) are not intentionally produced. They are formed and released unintentionally during thermal processes. PCDDs and PCDFs tended to be released by waste incinerators (Karasek and Dickson, 1987). The combination of elevated temperatures and the presence of chlorine from e.g. polyvinylchloride (PVC) led to the formation of the extremely toxic PCDDs and PCDFs. Stack emissions could contaminate entire areas around the incinerators with consequences for the quality of cow milk or local crop. Dioxins were first discovered after the Seveso (Italy) disaster (1976) when high quantities of dioxins were released after an explosion in a trichlorophenol factory (Mocarelli et al., 1991). Meanwhile, in many countries incinerators have been improved by changing the processes and installing appropriate filters.

Table 1. Stockholm Convention criteria for persistence, bioaccumulation, toxicity and long range transport of POPs.

Persistence	
i	Evidence that the half-life of the chemical in water is greater than two months, or that its half-life in soil is greater than six months, or that its half-life in sediment is greater than six months; or
ii	Evidence that the chemical is otherwise sufficiently persistent to justify its consideration within the scope of this Convention
Bioaccumulation	
i	Evidence that the bio-concentration factor or bio-accumulation factor in aquatic species for the chemical is greater than 5,000 or, in the absence of such data, that the log K_{ow} is greater than 5
ii	Evidence that a chemical presents other reasons for concern, such as high bioaccumulation in other species, high toxicity or ecotoxicity; or
iii	Monitoring data in biota indicating that the bio-accumulation potential of the chemical is sufficient to justify its consideration within the scope of this Convention
Long range transport potential	
i	Measured levels of the chemical in locations distant from the sources of its release that are of potential concern
ii	Monitoring data showing that long-range environmental transport of the chemical, with the potential for transfer to a receiving environment, may have occurred via air, water or migratory species; or
iii	Environmental fate properties and/or model results that demonstrate that the chemical has a potential for long-range environmental transport through air, water or migratory species, with the potential for transfer to a receiving environment in locations distant from the sources of its release. For a chemical that migrates significantly through the air, its half-life in air should be greater than two days
Adverse effects	
i	Evidence of adverse effects to human health or to the environment that justifies consideration of the chemical within the scope of this Convention; or
ii	Toxicity or ecotoxicity data that indicate the potential for damage to human health or to the environment

Structures and use

Whereas all initial POPs were chlorinated chemicals and mainly pesticides, POPs that were added at a later stage also included brominated and fluorinated compounds and chemicals with a more industrial application. Brominated diphenylethers (PBDEs) and hexabromocyclododecane (HBCD) belong to the group of brominated flame retardants. These chemicals are being produced in high quantities. Many national legislations require the use of flame retardants in many materials, such as electric and electronic systems (TV, cell phones, computers), furniture and building materials. Although the PBDEs and HBCD have been banned in most countries, other brominated flame retardants are still being produced in annually growing volumes.

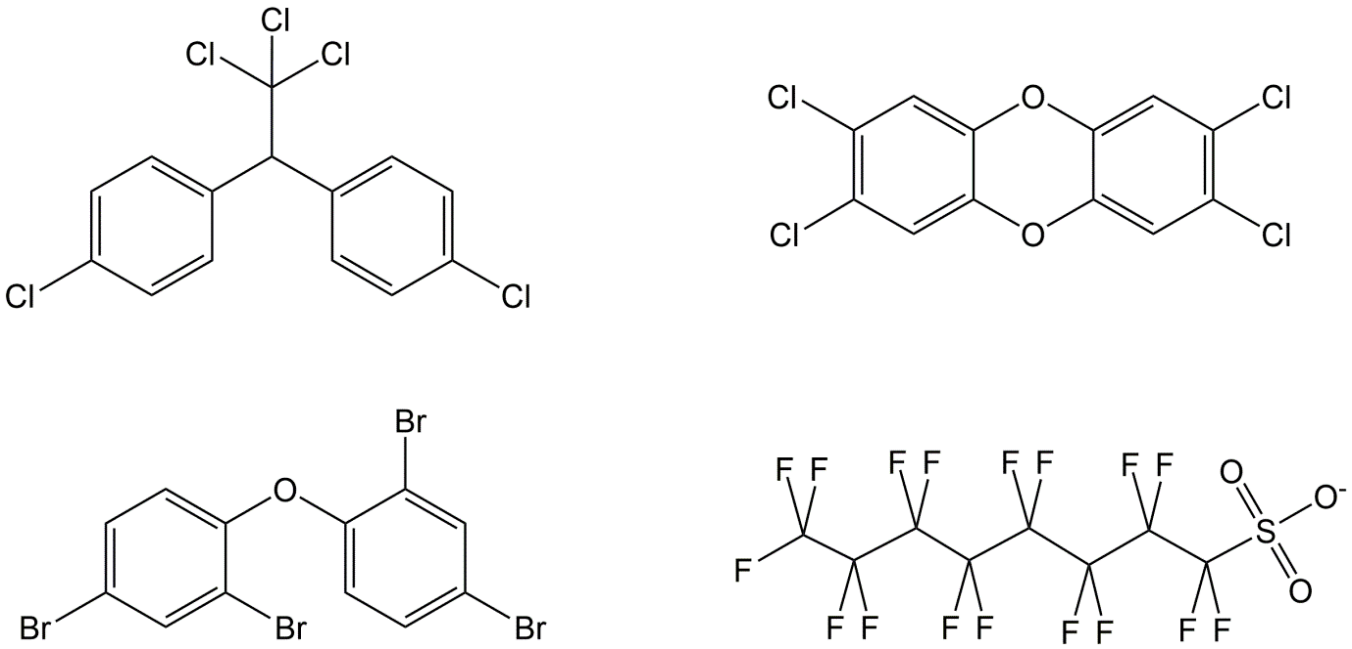


Figure 1. Structures of p,p'-DDT, 2,3,7,8-tetrachloro-p-dioxin, 2,4,2',4'-tetrabromodiphenylether (a specific PBDE) and perfluoroctane sulfonic acid (PFOS).

Perfluorinated alkyl substances (PFASs) have many applications. Examples are Teflon production, use in fire-fighting foams, ski wax, as dirt and water repellant on outdoor clothes and carpets and many more. They are different from most other POPs because they are both lipophilic and hydrophilic due to a polar group present in most of the molecules. Examples of structures of a few POPs are given in Figure 1.

Persistence and bioaccumulation

The carbon-halogen bond is so strong that any type of degradation is unlikely to occur or will only occur on the long term and to a minor extent. Due to the size of the halogen atom, the strength of the halogen-carbon bond decreases from C-F to C-Cl, C-Br and C-I. In addition, these halogenated chemicals are lipophilic and, therefore, easily migrate to lipids, such as in living organisms. Because fish is a primary target, POPs enter the food chain in this way and biomagnification can occur (De Boer et al., 1998). High levels of POPs are, consequently, found in marine mammals (seals, whales, polar bears) and also in humans (Meironyte, 1999). Women may transfer a part of their POP load again to their children, the highest quantities to their firstborns.

Long range transport

Chemicals that migrate significantly through the air with a half-life in air greater than two days qualify for the POP criterion of long range transport. Many chemicals are indeed transported by air, often in different stages. Chemicals are emitted from a stack or evaporate from the soil in relatively warm areas and travel in the atmosphere toward cooler areas, condensing out again when the temperature drops. This process, repeated in 'hops', can carry them thousands of kilometers within days. This is called the 'grasshopper effect' (Gouin et al., 2004). It results in colder climate zones, in particular countries around the North Pole, receiving relatively high amounts of POPs.

Adverse environmental and health effects

There is very little doubt on the toxicity of POPs. Of course, the dose is always determining if a compound is causing an effect in the environment or in humans. POPs, however, are very toxic at very low doses. The Seveso disaster showed the high toxicity of

dioxins for humans. Polybrominated biphenyls (PBBs) caused a high mortality in cattle when they were inadvertently fed with these chemicals (Fries and Kimbrough, 2008). Evidence of toxicity is often coming from laboratory studies with animals (*in vivo*) and more recently from *in vitro* studies. These studies are in particular important for the assessment of chronic toxicity. Many POPs are carcinogenic or act as endocrine disruptor.

Analysis

The analysis of POPs in environmental or human matrices is relatively complicated and costly. The compounds need to be isolated from the matrix by extraction. Subsequently, the extracts need to be cleaned from interfering compounds such as fat from organisms or sulphur in case of sediment or soil samples. Finally, due to the required sensitivity and selectivity, expensive instrumentation such as gas or liquid chromatography combined with mass spectrometry is needed for their analysis (Muir and Sverko, 2006). UN Environment is investing in large capacity building programs to train laboratories in developing countries in this type of analysis. According to the Stockholm Convention, countries shall manage stockpiles and wastes containing POPs in a manner protective of human health and the environment. POPs in wastes are not allowed to be reused or recycled. A global monitoring program has been installed to assess the effectiveness of the Convention.

Future

Much has to be done to achieve the original goals of eliminating the production and use of POPs and gradually reduce their spreading into the environment. A global treaty such as the Stockholm Convention with 182 countries involved is in a continuous challenge with procedures and political realities in countries, which hamper the achievement of perceived simple goals such as to eliminate the use of PCBs in 2025. The goals are, however, extremely important, as POPs are a global threat for current and future generations.

References

De Boer, J., Wester, P.G., Klamer, J.C., Lewis, W.E., Boon, J.P. (1998). Brominated flame retardants in sperm whales and other marine mammals - a new threat to ocean life? Nature 394, 28-29.

Fiedler, H., Kallenborn, R., de Boer, J., Sydnes, L.K. (2019). United Nations Environment Programme (UNEP): The Stockholm Convention - A Tool for the global regulation of persistent organic pollutants (POPs). Chem. Intern. 41, 4-11.

Fries, G.F., Kimbrough, R.D. (2008). The PBB episode in Michigan: An overall appraisal. CRC Critical Rev. Toxicol. 16, 105-156.

Gouin, T., Mackay, D., Jones, K.C., Harner, T., Meijer, S.N. (2004). Evidence for the "grasshopper" effect and fractionation during long-range atmospheric transport of organic contaminants. Environ. Pollut. 128, 139-148.

Karasek, F.W., Dickson, L.C. (1987). Model studies of polychlorinated dibenzo-p-dioxin formation during municipal refuse incineration. Science 237, 754-756.

Meironyte, D., Noren, K., Bergman, Å. (1999). Analysis of polybrominated diphenyl ethers in Swedish human milk. A time-related trend study, 1972-1997. J. Toxicol. Environ. Health Part A 58, 329-341.

Mocarelli, P., Needham, L.L., Marocchi, A., Patterson Jr., D.G., Brambilla, P., Gerthoux, P.M. (1991). Serum concentrations of 2,3,7,8-tetrachlorodibenzo-p-dioxin and test results from selected residents of Seveso, Italy. J. Toxicol. Environ. Health 32, 357-366.

Muir, D.C.G., Sverko, E. (2006). Analytical methods for PCBs and organochlorine pesticides in environmental monitoring and surveillance: a critical appraisal. Anal. Bioanal. Chem. 386, 769-789.

Van den Berg H. (2009). Global status of DDT and its alternatives for use in vector control to prevent disease. Environ. Health Perspect. 117:1656-63.

2.2.4. Question 1

Which criteria are used by the Stockholm Convention to define chemicals as POPs?

2.2.4. Question 2

What are the objectives of the Stockholm Convention on POPs?

2.2.4. Question 3

How can dioxins be formed?

2.2.4. Question 4

Why is the analysis of POPs in environmental and human matrices expensive?

2.2.5. Persistent Mobile Organic Chemicals (PMOCs)

Authors: Pim de Voogt

Reviewers: John Parsons, Hans Peter Arp

Leaning objectives:

You should be able to:

- define a substance's persistence
- define a partition coefficient
- understand the relationship between K_{OW}, D_{OW}, K_D and mobility
- understand the relationship between a substance's mobility and persistence on the one hand and its potential for human exposure on the other

Keywords: Mobility, persistence, PMT

Introduction

Ecosystems and humans are protected against exposure to hazardous substances in several ways. These include treating our wastewater so that substances are prevented from entering receiving surface waters, and purification of source waters intended for drinking water production.

Currently, a majority of the drinking water produced in Europe is either not treated or treated by conventional technologies. The latter remove substances by degradation (physical, microbiological) or by sorption. However, chemicals that are difficult to break down and that can pass through soil layers, water catchments and riverbanks and cross natural and technological barriers may eventually reach the tap water. Typically, these chemicals are persistent and mobile.

Polarity

When the electrons in a molecule are unevenly divided over its surface, this results in an asymmetric distribution of charge, with positive and negative regions. Such molecules have electric dipoles (see Figure 1) and are polar, in contrast to molecules where the charge is evenly distributed thus resulting in the molecule being neutral or apolar. The ultimate form of polarity is when a permanent charge is present in a compound. Such chemicals are called ionogenic. We distinguish between cations (having a permanent positive charge, e.g. protonated bases, and quaternary amines) and anions (negatively charged ions, e.g. dissociated acids, and organosulfates). Ionic charges in molecules can be pH dependent (e.g. acids and bases). Most, and in particular small, polar and ionic chemicals are water soluble, in other words they have a strong affinity to water (often referred to as hydrophilic). Because water is one of the most polar liquids possible (a strong negative charge on the oxygen and two strong positive charges on each hydrogen), this means that for very polar organic molecules solvation by water is more favorable energetically then sorption to solid particles.

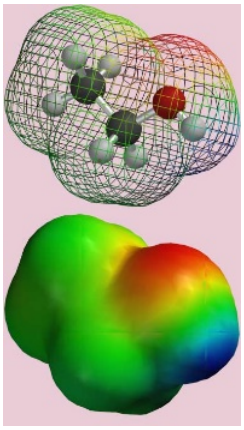


Figure 1. A model of ethanol showing the volume that is occupied by its electrons (A) and the direction of the dipole (B). The distribution of the electrons in an ethanol molecule is skewed relative to the protons to give a region having a partial negative charge, which is shown in red, and a corresponding region having a partial positive charge, which is shown in blue. (source: https://www.chem.uwec.edu/Chem150_Resources/content/elaborations/unitx/unit1-e-polarity.html and www.utdallas.edu/~biewerm/10-alcohols.pdf)

Chemicals that are nonpolar are inherently poorly water soluble and therefore tend to escape from the water compartment, resulting in their evaporation, or sorption to sediments and soils, or uptake and accumulation in organisms. It is therefore relatively easy to remove them from water during water treatment. In contrast, mobile organic chemicals, especially those that do not breakdown easily, pose a more serious threat for (drinking) water quality because they are much more difficult to remove. It should be noted that mobility and polarity can be thought of as a gradient, rather than a distinct category, with water being the most polar molecule, a large aliphatic wax being the most non-polar molecule, and all other organic molecules falling somewhere in the spectrum between.

In a recent study contaminants were analysed in Dutch water samples covering the journey from WWTP effluent to surface water to groundwater and then to drinking water. While the concentration level of total organic contaminants decreased by about 2 orders of magnitude from the WWTP effluents to the groundwater used for drinking water production, the hydrophilic contaminants (using chromatographic retention time as an indicator for hydrophilicity) in the WWTP effluents remained in the water throughout its passage to groundwater and into the drinking water (see Figure 2).

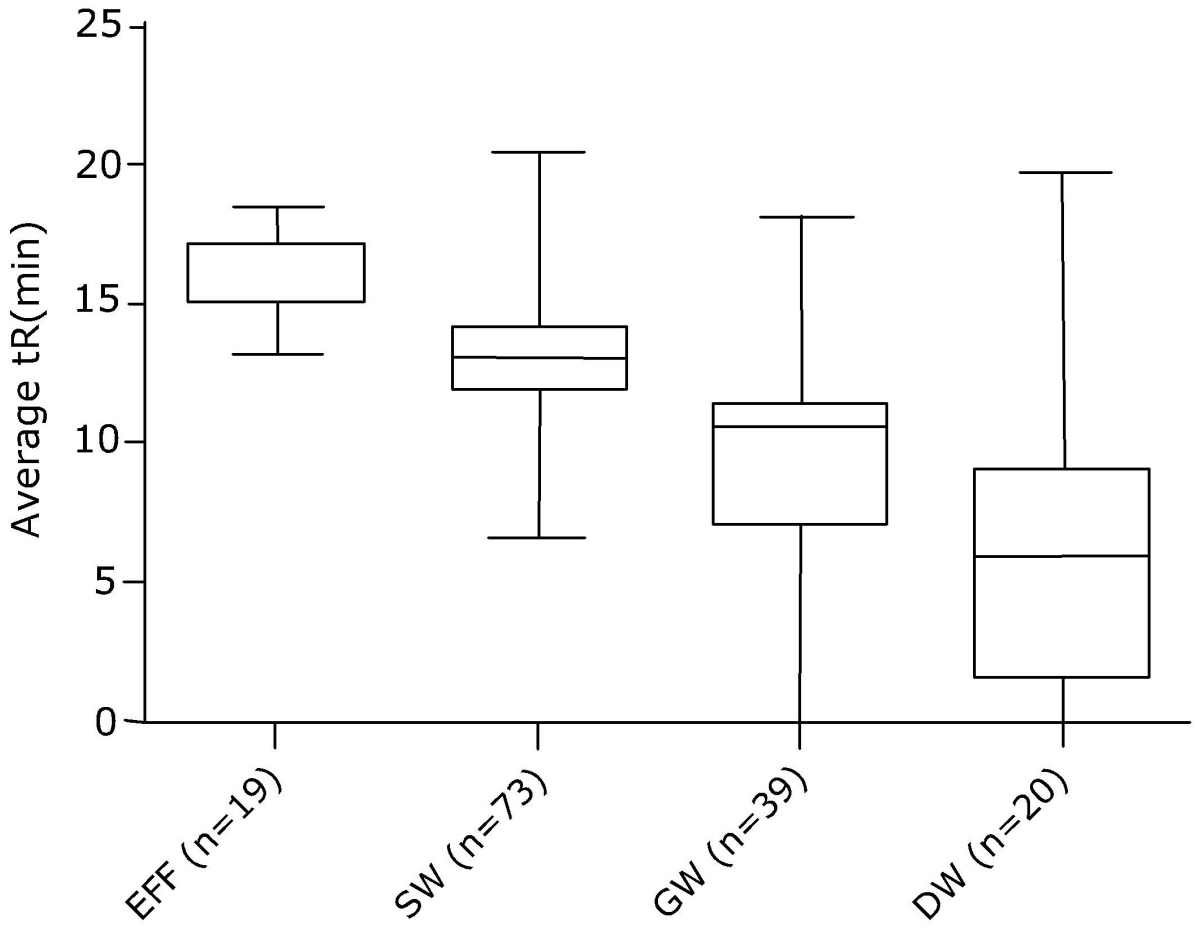


Figure 2. Average chromatographic retention time (tR) as a measure of average hydrophilicity of contaminants present in different water types. EFF, effluent; SW, surface water; GW, groundwater; DW, drinking water. Redrawn from Sjerps et al. (2016) by Wilma IJzerman.

Mobility and persistence

The mobility of chemicals in aquatic ecosystems is determined by their distribution between water and solid particles. The more the substance has an affinity for the solid phase the less mobile it will be. The distribution coefficient is known as $K_{D,}$ which expresses the ratio between the concentrations in the solid phase (soil, sediment, suspended particles), C_S, and the dissolved phase at equilibrium, C_W, i.e. $K_D = C_S/C_W$. For neutral non-polar chemicals the distribution is almost entirely determined by the amount of organic carbon in the solid phase, f_{OC}, and hence their distribution is usually expressed by K_{OC}, the organic carbon-normalized K_D (i.e. $K_{OC} = K_D/ f_{OC}$). Unfortunately, there are relatively few reliable K_D or K_{OC} data available, in particular for polar chemicals. Instead, K_{OW} is often used as a proxy for K_{OC}. The n-octanol/water partition coefficient: K_{OW}, is the equilibrium distribution coefficient of a chemical between n-octanol and water, $K_{OW} = C_{octanol}/C_{water}$. It's logarithmic value is often used as a proxy to express the polarity of a compound: a high log K_{OW} means that the compound favors being in the octanol phase rather than in water, which is typically the case for a nonpolar compound. For ionizable chemicals we need to account for the pH dependency of K_{OW}: at low pH an organic acid will become protonated (this in turn depends on its pK_a value) and thus less polar. D_{OW} is the pH-dependent K_{OW}. It can be assumed that ions, whether cationic or anionic, will no longer dissolve into octanol but rather be retained in the water, because ions have much higher affinity for water than for octanol.

Accounting for this, for organic acids, the pH dependency of the D_{OW} can be expressed as:



$$D_{OW} = \frac{K_{OW}}{1 + 10^{pH - pK_a}}$$

Therefore, as pH increases above the pKa, the smaller the D_{OW} will get in the case of organic acids. In the case of basis, the opposite is true; the more the pH dips below the pKa of an organic base, the more cations form and the lower the Dow becomes.

However, one has to keep in mind that the assumption that the (log) K_{OW} or D_{OW} value inversely correlates with a compound's aquatic mobility is, certainly, very simplistic. The behavior of an ionic solute will obviously also be determined by interactions i) with sites other than organic carbon, e.g. ionizable or ionic sites on soil and sediment particles, and ii) with other ions in solution.

The persistence of a compound is assessed in experimental tests by monitoring the rate of disappearance of the compound from the most relevant compartment. This is often done using standardized test protocols. In the European REACH legislation of chemicals, criteria have been established to qualify chemicals as persistent (P) and "very persistent"(vP) based on the outcomes of such tests. Table 1 presents the P and vP criteria used. Unfortunately, good-quality experimental data on half-lives are rare and obtaining such data requires both time-consuming and expensive testing.

Currently there is no certified definition of a compound's mobility (M). Several possible compound properties have been proposed to characterize mobility, including a compound's aqueous solubility or its K_{OC} value. If (experimental) K_{OC} values are not available, D_{OW} values can be used as a proxy.

Table 1. *P and vP criteria identical to Annex XIII to the REACH regulation (source: ECHA chapter R.11. Version 3.0, June 2017)*

	Persistent (P) in any of the following situations	**Very persistent (vP) in any of the following situations**
Freshwater	Half-life > 40 days	Half-life > 60 days
Marine water	Half-life > 60 days	Half-life > 60 days
Freshwater sediment	Half-life > 120 days	Half-life > 180 days
Marine sediment	Half-life > 180 days	Half-life > 180 days
Soil	Half-life > 120 days	Half-life > 180 days

Table 2. *Proposed cut off values of compound properties proposed by the German Environmental Agency (UBA) to define substance mobility**

	Mobile (M) if compound is P or vP and any of the following situations	**very Mobile (vM) if compound is P or vP and any of the following situations**
Lowest experimental log K_{OC} (at pH 4-9)	≤4.0	≤3.0
Log D_{OW} (at pH 4-9)	≤4.0 if no experimental Log K_{OC} data available	≤3.0 if no experimental Log K_{OC} data available

* note that the proposed criteria may change by the time of publication

Regulation and gaps in knowledge

The majority of chemicals for which international guidelines exist or that are identified as priority pollutants in existing regulations (e.g. EU Water Framework Directive and REACH), are nonpolar with log D_{OW} values mostly above two (see Figure 3b). The German Ministry of Environment (UBA) has recently proposed to develop regulation for chemicals with P, M and toxic (T) properties (PMT substances) analogous to the existing PBT criteria used for regulation of chemicals in the EU. UBA proposed to use cut-off values of the K_{OC} or D_{OW} (if K_{OC} data are not available) to define Mobile (M) or very mobile (vM) in conjunction with persistence criteria (see Table 2). Note that the K_{OC} and D_{OW} values have to be obtained from testing at an environmentally relevant pH range (pH 4-9).

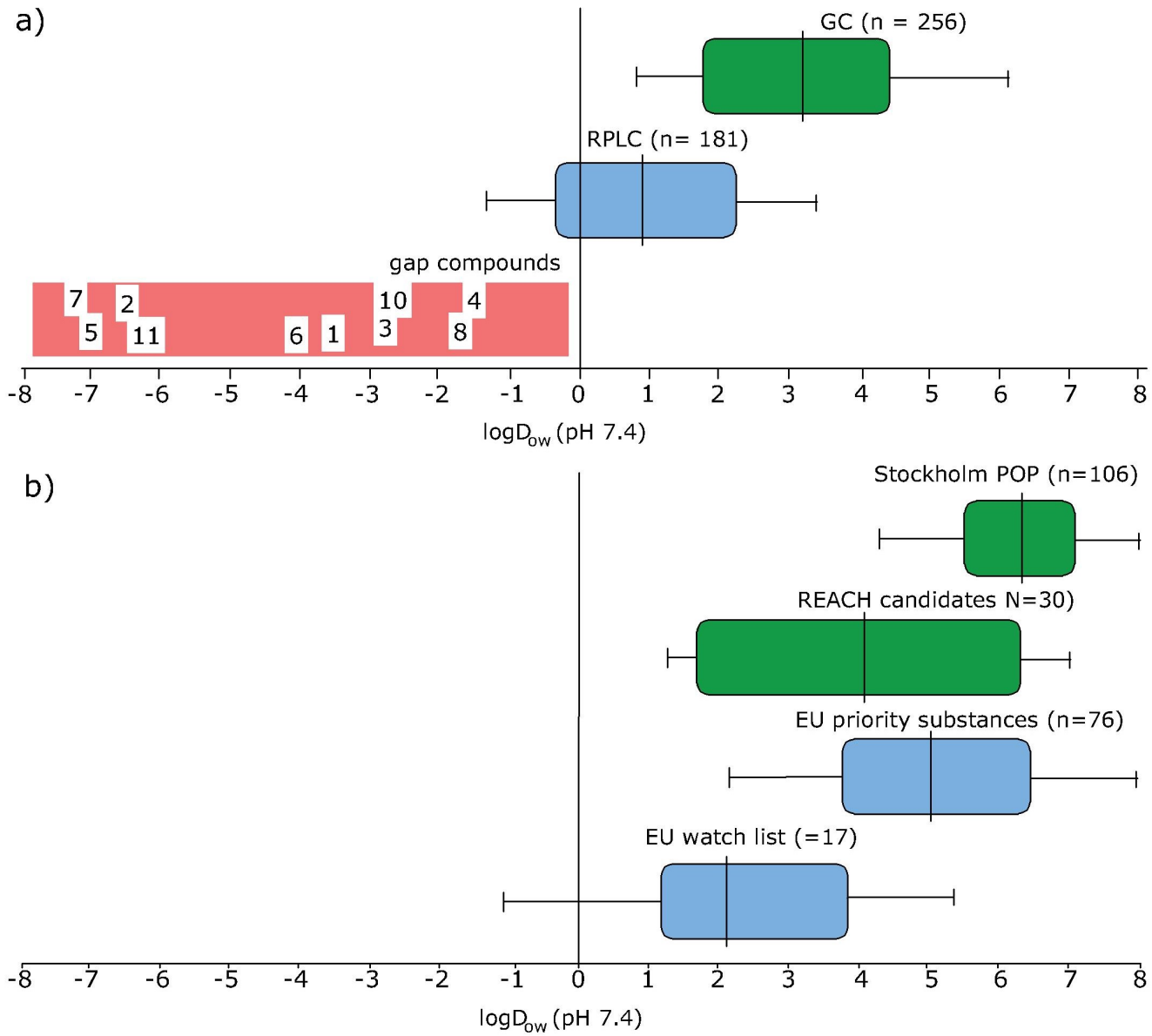


Figure 3. Box and whisker plots of calculated $logD_{OW}$ values at pH 7.4 of: (a) contaminants in water analyzed by either GC-MS or LC-MS and of examples of mobile chemicals; (b) contaminants regulated by the Stockholm Convention; candidates of Substances of Very High Concern (SVHCs) according to REACH, Article 57 d−f; the list of priority substances according to the Water Framework Directive (WFD); and the so-called Watch List of the WFD. The whiskers point to the 10th and 90th percentile. Numbers in (a) refer to 1: Aminomethylphosphonic acid (AMPA), 2: Paraquat, 3: Cyanuric acid, 4: N,N-dimethylsulfamide (DMS), 5: Diquat, 6: 5-Fluorouracil, 7: Glyphosate, 8: Melamine, 9: Metformin, 10: Perfluoroacetic acid, 11: EDTA. Redrawn from Reemtsma et al. (2016) by Wilma IJzerman.

When we consider current analytical techniques used for monitoring contaminants in the environment, it can be readily seen that the scope of techniques most often used (gas chromatography, GC, and reversed-phase liquid chromatography, RPLC) do not overlap with what is required for chemicals having log D_{OW} values typical of the most mobile chemicals, having a log Dow below zero (see Figure 3a). Consequently, there is limited information available on the occurrence and fate of these mobile chemicals in the environment. Nevertheless, some examples of persistent and mobile chemicals have been identified. These include highly polar

pesticides and their transformation products, for instance glyphosate and aminomethylphosphonic acid (AMPA), short-chain perfluorinated carboxylates and sulfonates, quaternary ammonium chemicals such as diquat and paraquat and complexing agents such as EDTA. There are, however, likely to be many more chemicals that could be classified as PMOC and we can therefore conclude that there is a gap in the knowledge and regulation of persistent mobile organic chemicals.

References

Arp, H.P.H., Brown, T.N., Berger, U., Hale, S.E. (2017). Ranking REACH registered neutral, ionizable and ionic organic chemicals based on their aquatic persistency and mobility. Environmental Science: Processes Impacts 19, 939-955.

Reemtsma, T., Berger, U., Arp, H. P. H., Gallard, H., Knepper, T. P., Neumann, M., Benito Quintana, J., de Voogt, P. (2016). Mind the gap: persistent and mobile organic chemicals - water contaminants that slip through. Environmental Science & Technology 50, 10308-10315.

Sjerps, R.M.A., Vughs, D., van Leerdam, J.A., ter Laak, T.L., van Wezel, A.P. (2016). Data-driven prioritization of chemicals for various water types using suspect screening LC-HRMS. Water Research 93, 254-264.

2.2.5. Question 1

What is the reason that PMOCs may end up in drinking water despite the application of removal processes?

2.2.5. Question 2

Explain why the average hydrophilicity of compounds present in drinking water is higher than that of surface waters or wastewater from which is has been produced.

2.2.6. Ionogenic organic chemicals

Authors: Steven Droge

Reviewer: John Parsons, Satoshi Endo

Leaning objectives:

You should be able to:

- understand that IOCs are abundantly present in many types of chemical classes, but all share common features such as dissociation of a proton from a polar moiety (acids to form anions) and association of a proton onto a polar moiety (bases to form cations).
- calculate the fraction of neutral/ionic species for each chemical at a given pK_a and pH.
- make some rough predictions about pK_a values from the IOC molecular structure.

Keywords: pK_a, dissociation constant, speciation, drugs, surfactants, solubility

Introduction

Ionogenic organic chemicals (IOCs) are widely used in industry and daily life, but also abundantly present as chemicals of emerging concern. For environmental risk assessment purposes, IOCs may be defined as organic acids, bases, and zwitterionic chemicals that under common environmental pH conditions exist for a large part as charged (ionic) species, with only modest fraction of neutral species. The environmentally relevant pH range could be argued to lie between 4 (acidic creeks, even lower in polluted streams from volcanic regions or mine drainage systems) to 9 (sewage treatment effluents). The environmental behaviour of IOC pollutants of concern are different from neutral chemicals of concern, because the aqueous pH controls the neutral fraction of dissolved IOCs and the ionic form is highly soluble and interacts partly via electrostatic interactions with environmental susbtrates. Note that the major fraction of an IOC can also be neutral in a certain environmental system, and in that case it is often the neutral form that dominates the chemical's behavior.

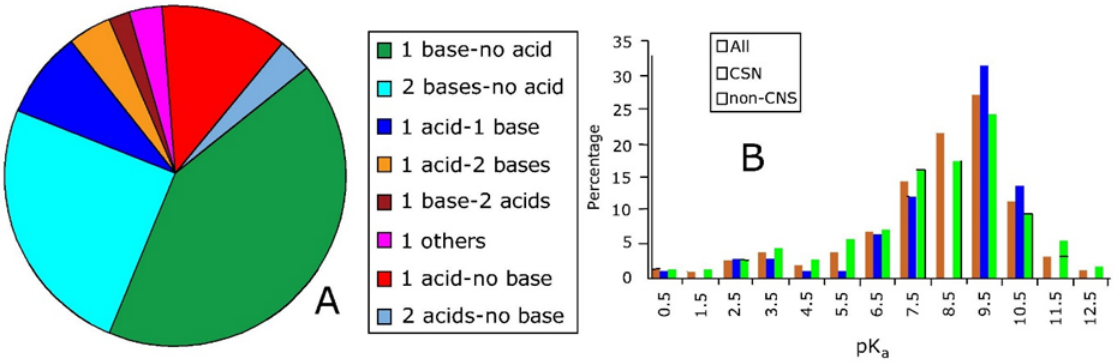


Figure 1. A survey on the 1999 World Drug Index (WDI) database revealed that >63% of the 51,596 chemicals had acidic or basic functionalities. Shown in the circle diagram A is the distribution of different ionizable groups in this fraction of 63% drugs, and in bar diagram B the distribution in pK_a values for basic drugs (CNS = drugs acting on central nervous system). Adapted from Manallack (2007) by Wilma IJzerman.

IOCs are common in many different types of pollutant classes. A random subset analysis of all EU (pre-)registered industrial chemicals indicated that large fractions of the total list of >100,000 chemicals are IOCs (51% neutral; 27% acids; 14% bases; 8% zwitterions/amphoterics). In another source, It has been estimated that >60% of all prescription drugs (Section 2.3.3) are IOCs (Manallack, 2007), and even higher fractions for illicit drugs (Section 2.3.4) (Figure 1). Well known examples are basic beta-blockers (e.g. propranolol), basic antidepressants (e.g. fluoxetine and sertraline), acidic non-steroidal anti-inflammatory drugs (NSAID such as diclofenac), and basic opioids (e.g. morphine, cocaine, heroin) and basic designer drugs (e.g. MDMA). The majority of surfactants and polyfluorinated chemicals (e.g. PFOS and GenX) are IOCs (Section 2.3.8), as well as wide variety of important pesticides (e.g. zwitterionic glyphosate) (Section 2.3.1) and (natural) toxins (Section 2.1) (e.g. peptide based multi-ionic cyanobacterial toxins).

Environmental behavior of IOCs

The release into the environment is specific for each of these types of IOCs with a different use, but in many cases happens via sewage treatment systems. If sorption to sewage sludge is very strong, application of sludge onto terrestrial (agricultural) systems is a key entry in many countries. However, many IOCs are rather hydrophilic and will mainly be present in wastewater effluent released into aquatic systems. As they are hydrophilic, they are considered rather mobile which allows for rapid transport through e.g. groundwater plumes, soil aquifers, and (drinking water) filtration steps. The distinction between (mostly) neutral chemicals and IOCs is important because the ionic molecular form generally behaves very differently from the corresponding neutral molecular form. For example, in many aspects ionic molecules are non-volatile compared to the corresponding neutral molecules, while neutral molecules are more hydrophobic than the corresponding ionic molecules. As a result of their lower "hydrophobicity", ionic molecules often bind with lower affinity to soils and are therefore more mobile in the environment. The ionic forms bioaccumulate to a lower extent and can therefore be less toxic than the corresponding neutral form (though not necessarily). However, there are various important exceptions for these rules. For example, clay minerals sorb cationic IOCs fairly strongly via ion exchange mechanisms. Certain proteins (e.g. the blood serum protein albumin) tightly bind anionic chemicals because of cationic subdomains on specific (enzymatic) pockets, which allows for effective transport throughout our systems and over cell membranes.

Calculating and predicting the dissociation constant (pK

The critical chemical parameter describing the ability to ionize is the acid dissociation constant (pK_a). The pK_a defines at which pH 50% of the IOC is in either the neutral or ionic form by releasing an H^+ from the neutral molecule acids (AH to anion A^-), or accepting an H^+ onto the neutral molecule base (B to cation BH^+). The equilibrium between neutral acid and dissociated form can thus be defined as:

$[HA] \leftrightarrow [H^+] + [A^-]$ (eq.1)

where the chemical's equilibrium speciation is defined as:

$$K_a = \frac{[H^+][A^-]}{[HA]}$$ (eq.2)

which gives the pK_a as:



$pK_a = -log(K_a)$ (eq.3)

and as a function of pH, the ratio of the acid and anion is defined as:

$pH = pK_a + log \frac{[A^-]}{[HA]}$ for acids and $pH = pK_a + log \frac{[B]}{[BH^+]}$ for bases (eq.4)

Although the term pK_b is also used to denote the base association constant, it is conventional we consider $[BH^+]$ as acid and use 'pK_a' and other relationships for bases as well. The fraction of neutral species (f_N) for simple IOCs (one acidic or basic site) can be readily calculated with a derivatization of the Henderson-Hasselbalch equation:

$f_N = \frac{1}{1 + 10^{\alpha(-pH+pK_a)}}$ (eq.5)

in which α = 1 for bases, and -1 for acids.

in which α = 1 for bases, and -1 for acids. Using equation 5, Figure 2 presents a typical speciation profile for an acid (shown with pKa 5, so perhaps a carboxylic acid) and a base (shown with pKa 9, so perhaps a beta-blocker drug). Following the curve of equation 5, it is interesting to see some simple rules: if the pH is 1 unit lower than the pKa, the deprotonated species fraction is present at 10%. If the pH is 2 units lower than the pKa, the deminished species fraction is present at 1%, 3 units lower would give 0.1%, etc. From this, it is easy to make a good estimate for the protonation of a strong basic drug like MDMA (pKa reported 9.9-10.4) in blood (pH 7.4): with a maximum of 3 units lower pH, up to 99.9% of the MDMA will be in the protonated form, and only 0.1% neutral. For toxicological modeling studies, e.g. in terms of permeation through the blood-brain barrier membrane, this is highly relevant.

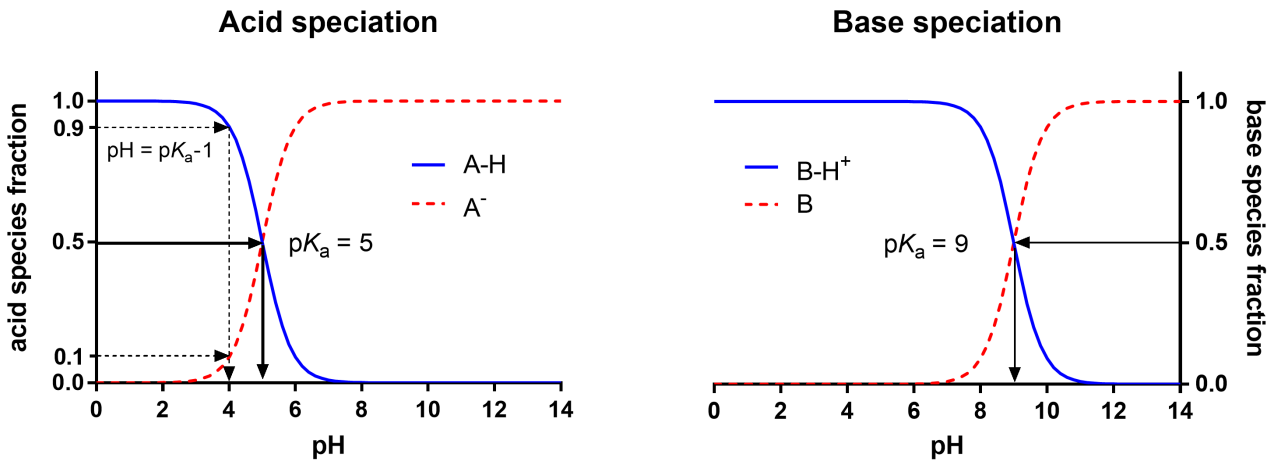


Figure 2. pH dependent speciation of an acid (with a pK_a of 5) and a base (with a pK_a of 9). As shown by the arrows, if pH = pKa, the IOCs exist 50% in the neutral form, and 50% in the ionic form. For the acid at pH = pKa -1 (pH 4), 90% is in the neutral form (AH), and 10% is in the negatively charged form (A^-). Drawn by Steven Droge.

Boxes 1-3. Extended learning: calculating the dissociation constant for multiprotic chemicals:
see end of this module

Acidic IOCs:

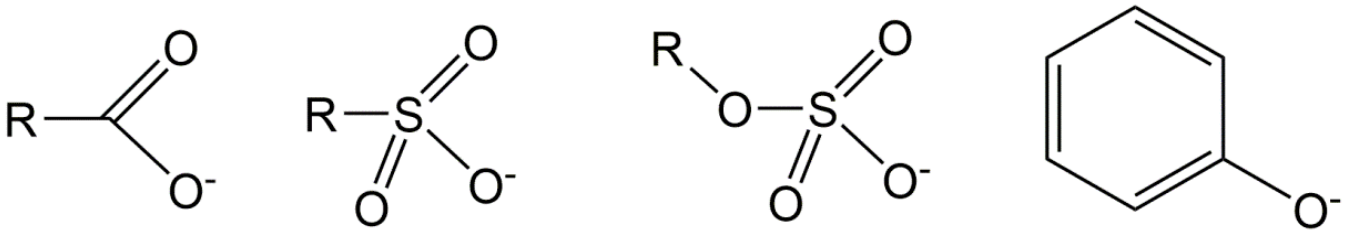


Figure 3. Different types of anionic moieties: carboxylate (anionic form of carboxylic acid), sulfonate (anionic form of sulfonic acid), sulfate (anionic form of sulfuric acid), phenolate (anionic form of phenol). (Source: Steven Droge)

For example, the painkiller (or non-steroidal anti-inflammatory drug, or NSAID) diclofenac is a carboxylic acid with a pK_a of 4.1. This means that at pH 4.1, 50% of the dissolved diclofenac is in the dissociated (anionic) form (so, $(1 - f_N)$ from equation 5). At pH 5.1 (1 unit higher than the pK_a) this is roughly 90% (90.91% to be more precise, but simply remembering 90% helps), at pH 6.1 (2 units higher than the pK_a) this is 99%. This stepwise increase in 50-90-99% with each pH unit works for all acids, and for bases the other way round. Test for yourself that at physiological pH of 7.4 (e.g. in blood) diclofenac is calculated to be 99.95% anionic.

Many carboxylic acids have a pK_a in the range of 4-5, but the neighboring molecular groups can affect the pK_a. Particularly electronegative atoms such as chlorine, fluorine, or oxygen may lower the pK_a, as they reduce the forces holding the dissociating proton to the oxygen atom. For example, trichloroacetic acid (CCl_3-COOH) has a pK_a of 0.77, while acetic acid (CH_3-COOH) has a pK_a of 4.75. For the same reason, perfluorinated carboxylic acids have a strongly reduced acidic pK_a compared to the analogous non-perfluorinated carboxylic acids.

Sulfate acids (see figure 3) are very strong acids, with a pK_a <0. These acids almost always occur in their pure form as a salt, for example the common soap ingredient sodium dodecylsulfate ("SDS" or "SLS", $Na.C_{12}H_{25}$-SO_4). Other common detergents are sulfonates, such as linear alkylbenzenesulfonate ("LAS", $C_{10\text{-}14}$-(benzyl-SO_3)), where the SO_3 anionic moiety is attached to a benzene ring, which can be positioned to different carbon atoms of a long alkyl chain. Even at the lower environmental pH range of about 4 these soap chemicals are fully in the anionic form. Such very strong acids, but also many weaker acids, are thus often sold in pure form as salts with sodium, potassium, or ammonium, which causes them to have different names and CAS numbers (e.g. $Na.C_{12}H_{25}$-SO_4 or $K.C_{12}H_{25}$-SO_4) than the neutral form.

Many phenols have a pK_a > 8, and are therefore mostly neutral under environmental pH. Electron-withdrawing groups on the aromatic ring of the phenol group, such as Cl, Br and I, can lower the dissociation constant. For example, the dinitrophenol-based pesticide dinoseb has a pK_a of 4.6, and is thus mostly anionic in the aquatic environment. Note that a hydroxyl group (-OH) not connected to an aromatic ring, such as -OH of alcohols, can in most cases for risk assessment be considered permanently neutral.

To help interpret the differences in pK_a between molecules, it sometimes helps to remember that in more acidic solutions, there are simply higher H^+ concentrations, in a logarithmic manner on the pH scale. At pH 3, the concentration H^+ protons in solution is 1 mM, while at pH 9 the H^+ concentration is 1 nM (6 pH units equals 10^6 times lower concentrations). The affinity ("K_a") of H^+ to associate with a negatively charged molecular group, is so low for strong acids that even at very high dissolved H^+ concentrations (low pH) only very few AH bonds (neutral acid fraction) are actually formed. In other words, for chemicals with a low K_a, even at low pH the neutral fraction is still low. For weak acids such as phenols, already at a very low dissolved H^+ concentrations (high pH) many AH bonds (neutral acid fraction) are formed. So it can be reasoned that the affinity of common acidic groups to hold on to a proton is in the order:

SO_4 < SO_3 < CO_2 < amide (C(=O)NH) < phenol < hydroxyl

Basic IOCs:

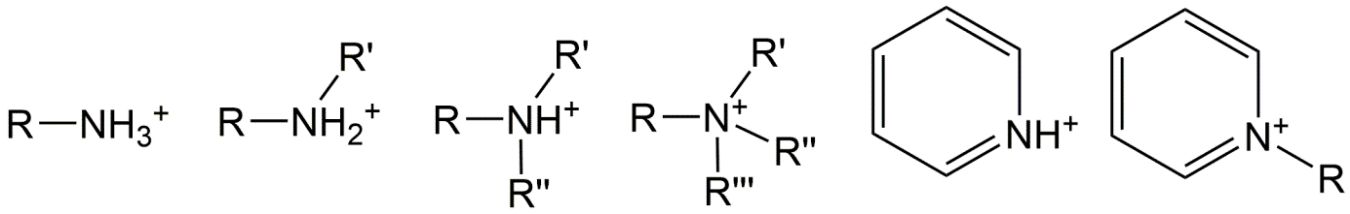


Figure 4. Different types of cationic moieties: primary amine, secondary amine, tertiary amine, quaternary ammonium (permanently charged), pyridinium (cationic form of pyridine), alkylpyridinium (permanently charged). (Source: Steven Droge).

For bases, it is mostly a nitrogen atom that can accept a proton to form an organic cation, because of the lone electron pair in nitrogen. Neutral nitrogen atoms have opportunities for 3 bonds. A primary amine group has the nitrogen atom bonding to only 1 carbon atom (represented here as part of a molecular fragment R), and two additional bonds with hydrogen atoms. The remaining electron pair readily accepts another proton to form a cationic molecule [$R\text{-}NH_3^+$]. Neutral secondary amines have one bond to hydrogen and two bonds to carbon atoms and can accept a proton to form [$R\text{-}NH_2^+\text{-}R'$], whereas neutral tertiary amines have no bonds to hydrogen but only to carbon atoms and can form [$R\text{-}NH^+\text{-}(R')(R'')$]. Of course, each R group may be the same (e.g. a methyl unit).

Many basic chemicals have complex functionalities that can influence the pK_a of the nitrogen moiety. However, as shown in the examples of figure 5, as long as there are at least two carbon atoms between the amine and the polar molecular fragment (for example OH, but much stronger for =O), the pK_a of the basic nitrogen groups in all three types of bases (primary, secondary and tertiary amines) is high, often above 9 (dissolved H^+ concentration $<10^{-9}$). So even at very low H^+ concentrations, dissolved protons like to be associated to such amine groups. As a result, amines such as most beta-blockers and amphetamine based drugs are predominantly positively charged molecules (organic cationic amines) in the common environmental pH range of 4-9, as well as in the pH of most biotic tissues that are useful for toxicological assessments. As soon as a polar group with oxygen (e.g. ketones or hydroxyl groups) is connected to the second carbon away from the nitrogen (e.g. $R\text{-}CH(OH)\text{-}CH_2\text{-}NH_2$) the pK_a is considerably lowered. Also nitrogen atoms as part of an aromatic ring, or connected to an aromatic ring, have much lower pK_a's: protons have rather low affinity to bind to these N-atoms and only start doing so if the proton concentration becomes relatively high (solution becomes more acidic).

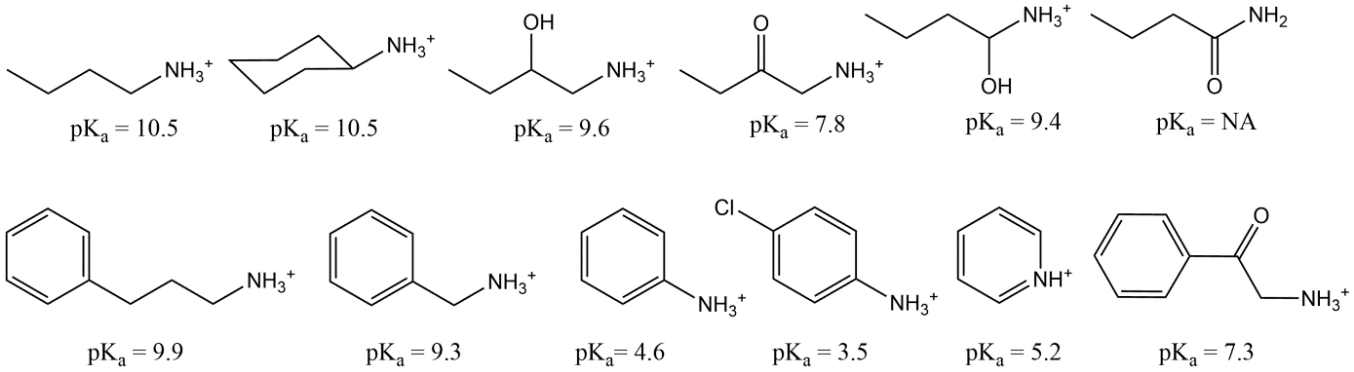


Figure 5. Different proton dissociation constants for amine groups: pKa is influenced by other functional structures. (Source: Steven Droge)

Relevance of accounting for electrostatic interactions

Most classical pollutants, such as DDT, PCBs and dioxins, are neutral hydrophobic chemicals. On the other hand, most metals are almost always cationic species (e.g. Cd^{2+}). Consequently, their environmental distribution and biological exposure are influenced by quite distinct processes. Obviously, predominantly charged IOCs behave somewhat in between these two extremes. The charged positive or negative groups cause a strong effect of electrostatic interactions between the IOC and environmental substrates (sorption or ligand/receptor binding). While also metals speciate into different forms, pH difference between environmental compartments can strongly influence the IOCs chemical fate and effect if ionizable group is relatively weak. An important difference to metals is that the nonionic molecular part still influences the IOC's hydrophobicity even in charged state for several processes.

As will be discussed in other chapters regarding chemical processes (see Chapter 3), it needs to be taken into account for IOCs that many environmental substrates (DOC, soil organic matter, clay minerals) are mostly negatively charged in the common range of environmental pH, but also that proteins involved in biotic uptake-distribution-effects are rich in ionogenic peptides that are part of binding pockets and reactive centers.

References

Manallack, D.T. (2007). The pKa distribution of drugs: application to drug discovery. Perspectives in Medical Chemistry 1, 25-38.

Box 1. Extended learning: calculating the dissociation constant for multiprotic chemicals:

Several common inorganic acids are multiprotic: they have multiple protons that can dissociate.

Multiples species can occur at a certain pH, such as for phosphoric acid (H_3PO_4, $H_2PO_4^-$, HPO_4^{2-} , HPO_4^{2-}), and carbonic acid (H_2CO_3, HCO_3^-, CO_3^{2-}). It is important to realize that there are actually two micro-species of HCO_3^-, because two hydroxylgroups can dissociate: HO-C(=O)-OH

A polyprotc acid H_nA can undergo *n* dissociations to form *n*+1 species. Each dissociation has a pK_a.

But how to calculate the fraction of each species of multiprotic chemicals?

The charge of a polyprotic acid can be described as $H_{n-j}A^{j-}$. A useful variable, *v*, can be defined for each general polyprotic acid:

$$v = \frac{(H_nA]}{[H^+]^n} \; so \; \frac{([H_nA])}{v} = [H^+]^n$$

for each dissociation reaction:

$H_{n-j+1}A^{-j+1} \leftrightarrow H^+ + H_{n-j}A^{j-}$

the dissociation constant K_j is:

$$K_j = \frac{[H^+][H_{n-j}A^{j-}]}{[H_{n-j+1}A^{-j+1}]}, and \; pK_j = -log(K_j)$$

The degree of dissociation of the acid (η) is equal to the ratio of the total charge (TC) to the total mol acid (TM). For a diprotic acid, a plot of η as a function of pH provides the dissociation curve:

$$\eta = \frac{TC}{v}\frac{v}{TM} = \frac{TC}{TM} = \frac{1K_1[H^+] + 2K_1K_2}{[H^+]^2 + K_1[H^+] + K_1K_2}$$

which can be a fitted curve to experimental data.

The degree of protolysis for the j^{th} species, α_j, can be calculated from the ratio of $[H_{n-j}A^{j-}]$/TM

$$\alpha_0 = \frac{[H_2A]}{v}\frac{v}{TM} = \frac{[H^+]^2}{[H^+]^2 + K_1[H^+] + K_1K_2}$$

$$\alpha_1 = \frac{[HA^-]}{v}\frac{v}{TM} = \frac{K_1[H^+]}{[H^+]^2 + K_1[H^+] + K_1K_2}$$

$$\alpha_2 = \frac{[A^{2-}]}{v}\frac{v}{TM} = \frac{K_1K_2]}{[H^+]^2 + K_1[H^+] + K_1K_2}$$

You can set up such a calculation in MS Excel, with calculations of α_0, α_1, α_2, at a range of different pH values ($[H^+]$ concentrations), for a given K_1 and K_2, and plot the speciation against pH.

More details are described by King *et al.* (1990) J. Chem. Educ. 67 (11), p. 932; DOI: 10.1021/ed067p932

Box 2. Example 1 for multiprotic chemicals: Carbonic acid

Let's try carbonic acid (bicarbonate, or H_2CO_3) as an example first. H_2CO_3 is the product of carbon dioxide dissolved in water. In pure water/seawater the hydration equilibrium constant $K_h = [H_2CO_3]/[CO_2] \approx 1.7\times10^{-3}$ / $\approx 1.2\times10^{-3}$ respectively, indicating that only 0.1% of dissolved CO_2 equilibrates to H_2CO_3. The dissolved concentration of CO_2 depends on atmospheric CO_2 levels according to the air-water distribution coefficient (Henry constant $k_H = pCO_2/[CO_2]$= 29.76 atm/(mol/L)). Because of the relevance of CO_2 in e.g. ocean acidification, and gas exchange in our lungs, it is interesting to see how H_2CO_3 speciates depending on pH. As in the formula H_nA, n= 2 for H_2CO_3.

$$v = \frac{([H_2CO_3])}{[H^+]^2}, \; so \; \frac{([H_2CO_3])}{v} = [H^+]^2$$

With pK_1^*=6.5 (in equilibrium with atmospheric CO_2) and pK_2=10.33, $K_1=10^{-6.5}$ and $K_2=10^{-10.3}$.

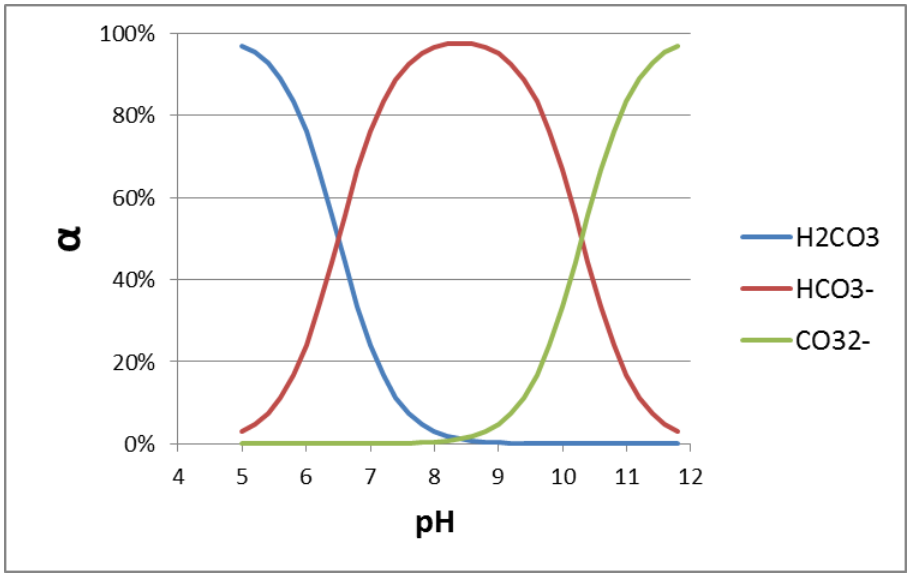
At pH 7, $[H^+]=10^{-7}$, so at pH 7 with these dissociation constants

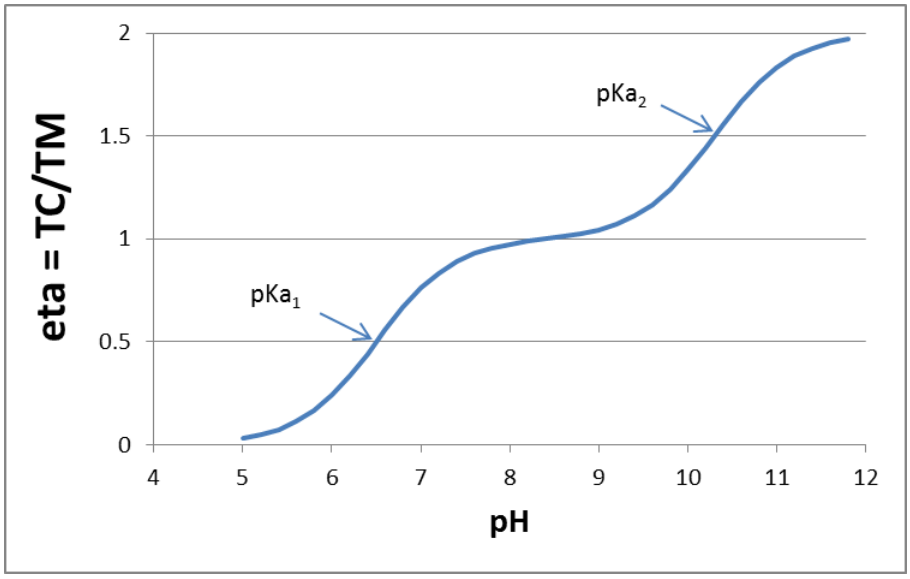
$$\alpha_0 = \frac{([H_2CO_3])}{v}\frac{v}{TM} = \frac{(10^{-7})^2}{(10^{-7})^2 + (10^{-6.5})(10^{-7}) + (10^{-6.5})(10^{-10.3})} = 0.24, so \; f_{H_2CO_3} = 24\%$$

$$\alpha_1 = \frac{[HCO_3^-]}{v}\frac{v}{TM} = \frac{(10^{-6.5})(10^{-7})}{(10^{-7})^2 + (10^{-6.5})(10^{-7}) + (10^{-6.5})(10^{-10.3})} = 0.76, so \; f_{HCO_3^-} = 76\%$$

$$\alpha_2 = \frac{[CO_3^{2-}]}{v}\frac{v}{TM} = \frac{(10^{-6.5})(10^{-10.3})}{(10^{-7})^2 + (10^{-6.5})(10^{-7}) + (10^{-6.5})(10^{-10.3})} = 0.00038, so f_{CO_3^{2-}} \sim 0\%$$

or from a series of Excel calculations at different pH:





you can copy/paste the following cells into cell A1 of a new Excel sheet, and extend the range of pH:

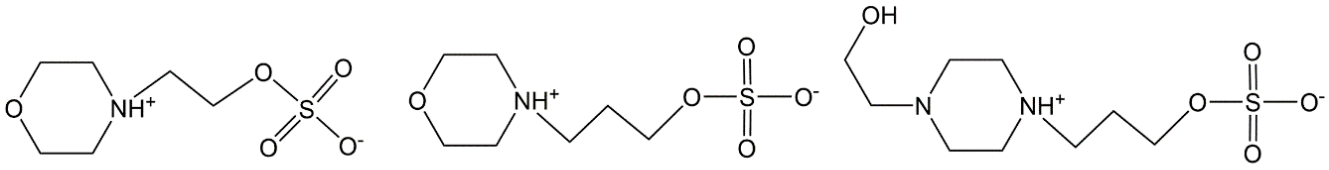
K1	3.16E-07
K2	5.01E-11
pH	5
[H+]	=10^-B3
a0 = H_2CO_3	=(B4*B4)/((B4*B4)+(B4*B1)+(B1*B2))
a1 = HCO_3^-	=(B4*B1)/((B4*B4)+(B4*B1)+(B1*B2))
a2 = CO_3^{2-}	=(B1*B2)/((B4*B4)+(B4*B1)+(B1*B2))
ETA	=((B4*B1)+2*(B1*B2))/((B4*B4)+(B4*B1)+(B1*B2))

Box 3. Example 2 for multiprotic chemicals: Zwitterions

Many organic pH buffers are zwitterionic chemicals, that contain both an acidic and a basic group. Norman Good and colleagues described a set of 20 of such buffers for biochemical and biological research (see for example www.interchim.fr/ft/0/062000.pdf, or www.applichem.com/fileadmin/Broschueren/BioBuffer.pdf). Examples are MES, MOPS, HEPPS (**Figure A**). These buffers are selected to:

- have a buffering pKa in the range of pH6-8 where most biochemical tests are performed;
- be readily soluble in water,
- be stable in test solutions, so resistant to (non)enzymatic degradation, not forming precipitates with salts
- ideally be impermeable to cell membranes so that they don't accumulate or reach active intracellular sites
- readily available, reasonably cheap.

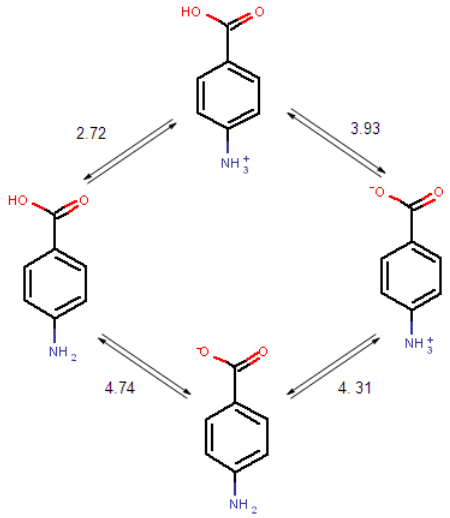
the zwitterionic chemicals with sulfate groups are actually always have the sulfate group charged, making it highly soluble and impermeable to cell membranes, while the amine group protonates between pH6-10, depending on neighbouring functional groups. The speciation of the amine groups in MES and MOPS simply follows the single pKa calculation of equations 1-5. In HEPPS, either of the two amines is protonated, the second pKa is 3, so the doubly charged molecules only occurs at much lower pH, but can still be used as a buffer.



MES pKa 6.1 MOPS pKa 7.2 HEPPS pKa 8.0

Figure A. MES, MOPS and HEPPS in charged form.

A zwitterionic chemical with two apparent pKa values relatively close is p-amino-benzoic acid. If chemicals have not one ionisable group, but *N* ionisable groups that speciation in a relevant pH range, than the amount of possible species is 2^N. So the zwitterion p-amino benzoic acid has 4 species, each with a separate pKa (pH where both species are present in equal concentrations).



Let's formulate the benzyl group in p-amino benzoic acid as X, the neutral amine base as B, and the neutral carboxylic acid as AH, so that the fully neutral species is BXAH.

Compared to the carboxylic acid, we now have under the most acidic conditions $(BHXAH)^+$, as α_0, the neutral species BXAH and the zwitterionic intermediate $(BHXA)^0$, as α_1, and anionic species at most alkaline conditions $(BXA)^-$, as α_2. The calculation of the fraction of each species can be calculated according to similar rules as for carbonic acid if the two dissociation constants are known (K1 = 2.4, K2=4.88).

However, this does not inform us on the ratio between the zwitterionic form and the fully neutral form. To do this, the speciation constants of the 4 microspecies are required.

[BH+XAH] ↔ [BXAH] + [H+] for which the pk1 is calculated to be 2.72

K_1=10^-2,72 = [BXAH]*[H+]/ [BH+XAH]

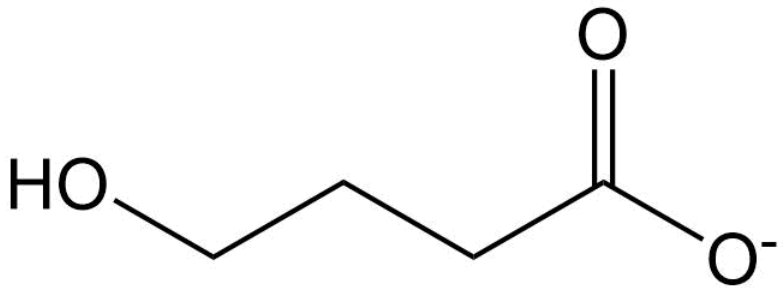
[BH+XAH]*10^-2.72 = [BXAH]*[H+]

which rearranges to [BXAH]= 10^-2.72 *[BH+XAH] /[H+]

[BH+XAH] ↔ [BH+XA-] + [H+] for which the pk2 is calculated to be 3.93
K_2=10^-3.93 = [BH+XA-]*[H+] / [BH+XAH]
[BH+XAH]*10^-3.93 = [BH+XA-]*[H+]
which rearranges to [BH+XA-] =10^-3.93*[BH+XAH]/[H+]
[BXAH] ↔ [BXA-] + [H+] for which the pk3 is calculated to be 4.74
K_3= 10^-4,74 = ([BXA-] * [H+]) / [BXAH]
[BXAH]*10^-4,74 = [BXA-] * [H+])
[BH+XA-] ↔ [BXA-] + [H+] for which the pk4 is calculated to be 4.31
K_4=10^-4,31 = [BXA-]*[H+] / [BH+XA-]
[BH+XA-]*10^-4,31 = [BXA-]*[H+]
So the ratio between zwitterionic form [BH+XA-] and neutral form [BXAH] equals to:
[BH+XA-] / [BXAH] = 10^-pK2 / 10^-pK1
[BH+XA-] / [BXAH] = 10^-3.93 / 10^-2.72 = 0.06: so only 6% zwitterionic vs 94% neutral species.
The macro pK1 and pK2 are then calculated as:
K1 = ([BXAH] *[H+] + [BH+XA-]*[H+])/ [BH+AXH]
K1 = [BXAH] *[H+]/[BH+XAH] + [BH+XA-]*[H+] / [BH+XAH]
K1 = 10^-pK1 + 10^- pK2
K1 = 10^-2.72 + 10^-3.93 =10^-2.69
K1=10^-2.69
K2 = ([BXA-]) *[H+] / ([BH+XA-] + [BXAH])
1/K2 = [BH+XA-]/([BXA-]) *[H+] + [BXAH] /([BXA-]) *[H+]
1/K2 = 1/10^-pK3 + 1/10^-pK4
1/K2 = 1/10^-4,31 + 1/10^-4,74 = 1/10^-4.87
K2 =10^-4.87

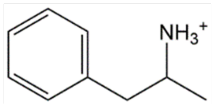
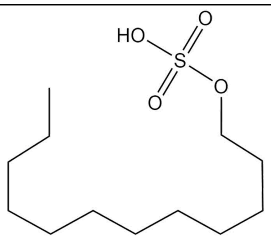
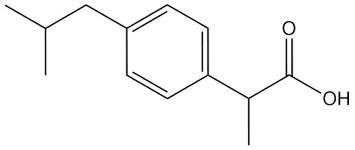
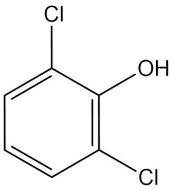
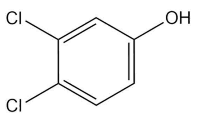
2.2.6. Question 1

Calculcate the neutral fraction at pH7.4 for the illicit drug GHB (apply HH equation of eq 5):



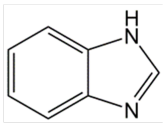
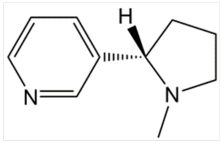
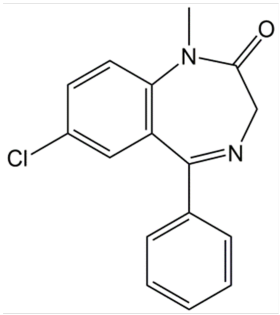
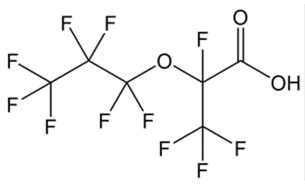
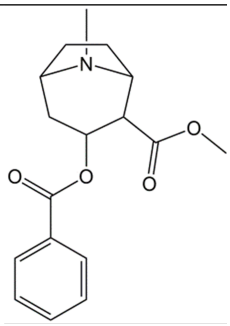
2.2.6. Question 2

Connect the chemicals to their appropriate pK_a values:

				
-2.0	5.2	6.8	9.9	8.6

2.2.6. Question 3

Decide/look up whether these chemicals are predominantly neutral/anionic/cationic in the environment:

benzimidazole	nicotine	diazepam	genX	cocaine
				

2.2.7. Complex mixtures/UVCBs

In preparation

2.2.8. Plastics

(draft)

Author: Ansje Löhr

Reviewer: John Parsons

Leaning objectives:

You should be able to:

- indicate the relevance of plastic for society
- describe the main characteristics of (micro)plastics
- describe the main ecological effects of (micro)plastics

Keywords: Plastic types, sources of plastics, primary and secondary microplastics, plastic degradation, effects of plastics

Introduction

Since its introduction in the 1950s, the amount of plastics in the environment has increased dramatically (Figure 1). A recent study by Jambeck *et al*. (2015) estimated that 192 coastal countries generated 275 million metric tonnes of plastic waste in 2010 of which around 8 million tons of land-based plastic waste ends up in the ocean every year. By UN Environment plastic pollution is seen as one of the largest environmental threats. If waste management does not change rapidly, another 33 billion tonnes of plastic will have accumulated around the planet by 2050. (Micro)plastics is widely recognized as a serious problem in the ocean, however, plastic pollution is also seen in terrestrial and freshwater systems.

Classification by size and morphology

Plastics are commonly divided into macroplastics and microplastics; the latter plastic particles are <5 mm in diameter (including nanoplastics). There are several ways to classify microplastics but the following two types are often used; primary microplastics and secondary microplastics. Primary microplastics have been made intentionally, like pellets or microbeads, secondary microplastics are fragmented parts of larger objects. Microplastics show a large variety in characteristics such as size, composition, weight, shape and color. These characteristics have an influence on the behaviour in the environment, like for instance, the dispersion in water and the uptake by organisms (Figure 2). Low-density particles float on water and are therefore more prone to advection than particles with a

higher density. Similarly, spheres are more likely to be taken up by organisms than fibers. The characteristics also affect the absorption of contaminants, adsorption of microbes, and potential toxicity.

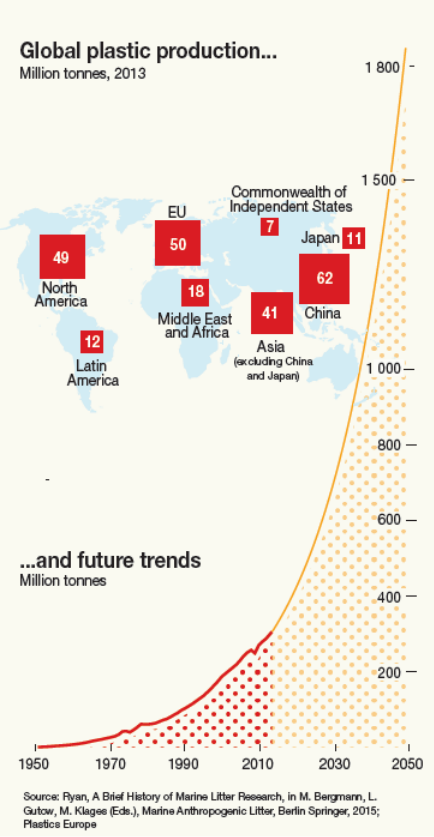


Figure 1. Global plastic production and future trends (source: http://www.grida.no/resources/6923; 2019, GRID-Arendal & Maphoto/Riccardo Pravettoni).

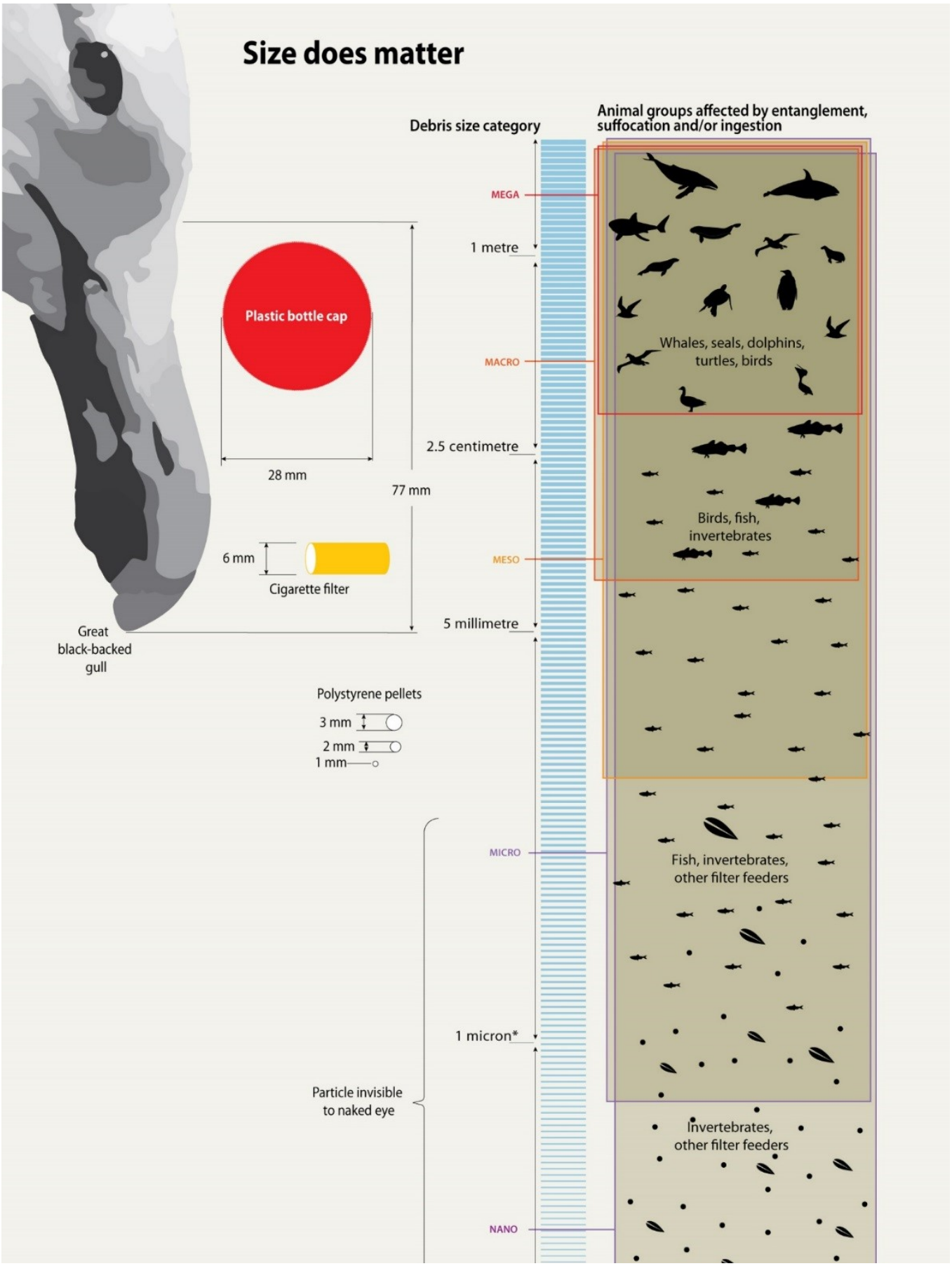

Size does matter
Debris size category
Animal groups affected by entanglement, suffocation and/or ingestion
MEGA
1 metre
MACRO
Whales, seals, dolphins, turtles, birds
2.5 centimetre
Birds, fish, invertebrates
MESO
5 millimetre
MICRO
Fish, invertebrates, other filter feeders
1 micron*
NANO
Invertebrates, other filter feeders
Particle invisible to naked eye
Plastic bottle cap
28 mm
77 mm
6 mm
Cigarette filter
Great black-backed gull
Polystyrene pellets
3 mm
2 mm
1 mm

Figure 2. Marine litter comes in all sizes. Large objects may be tens of metres in length, such as pieces of wrecked vessels, lost. (Source: http://www.grida.no/resources/6924; 2019; GRID-Arendal & Maphoto/Riccardo Pravettoni).

Classification by chemistry

Plastic is the term used to define a sub-category of the larger class of materials called polymers, usually synthesized from fossil fuels, although biomass and plastic waste can also be used as feedstock. Polymers are very large molecules that have characteristically long chain-like molecular architecture. There are many different types of plastics but the market is dominated by 6 classes of polymers: polyethylene (PE, high and low density), polypropylene (PP), polyvinyl chloride (PVC), polystyrene (PS, including expanded EPS), polyurethane (PUR) and polyethylene terephthalate (PET) (figure 3). In order to make materials flexible, transparent, durable, less flammable and long-lived, additives to polymers are used such as flame retardants (e.g. polybrominated diphenyl ethers), and plasticisers (e.g. phthalates). Some of these substances are known to be toxic to marine organisms and to humans.

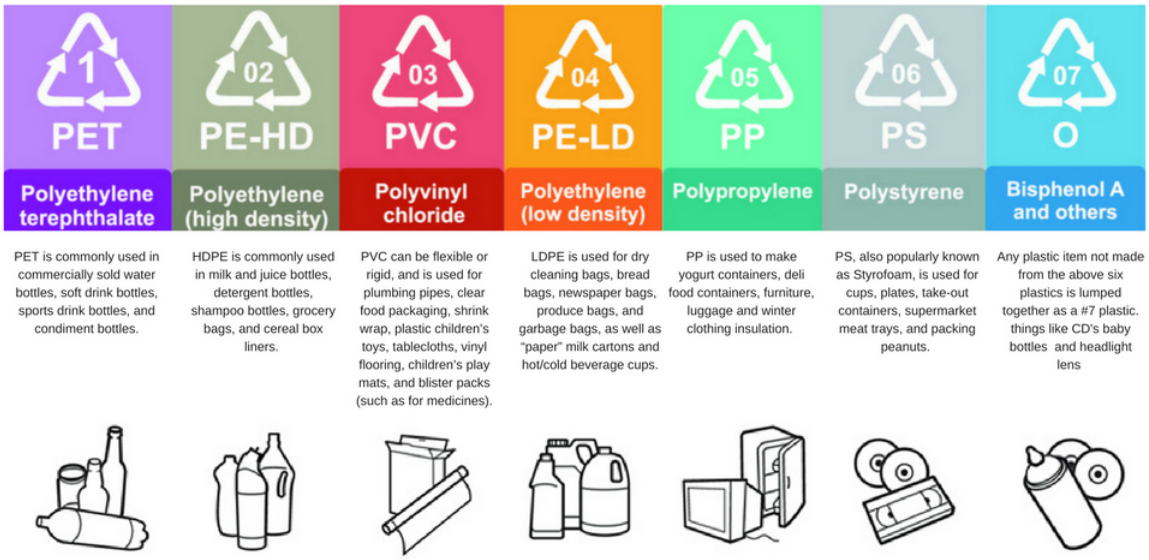


Figure 3.

Different types of plastics (source https://isustainrecycling.com/plastics-recycling/)

Biopolymers/ bioplastics

There is a lot of discussion on bioplastics as degradable plastics that these may still persist for a long time under marine conditions. Please watch this video by dr. Peter Kershaw.

Plastic degradation

Degradation of plastics takes place as soon as the plastic loses its original integrity and properties. There is a faster breaking up phase (degradation into microparticles) and a much slower mineralization phase (polymer chains being degraded to carbon dioxide). The degradation rate of plastics is determined by its polymer type, additive composition and environmental factors. Many commonly-used polymers are extremely resistant to biodegradation. Although plastics degrade in natural environments it is argued that no polymer can be efficiently biodegraded in a landfill site. Plastics in aquatic environments can be subject to in-situ degradation, e.g. by photodegradation or mechanical fragmentation but are in general very durable. As a result, plastics that are present in our oceans will degrade at a very slow pace, (Figure 4). So the majority of plastics produced today will persist in the environment for decades and probably for centuries, if not millennia.



Figure 4. : the time required to degrade different materials. (source https://futurism.com/plastic-decomposition)

Plastics in the environment

- *Sources and pathways:*

Plastics are found in terrestrial, freshwater, estuarine, coastal and marine environments, and even in very remote areas of the world and the deep-sea. Sources and pathways of marine litter are diverse and exact quantities and routes are not fully known. But there is a surge in interest to determine the exact quantities and types of plastic litter and pathways in the environment and most of the plastic in our oceans originates from land-based sources (Figure 5) but also from sea-based sources. Most PE and PP is used in (single-use) packaging products that have a short lifetime and end up soon as waste.

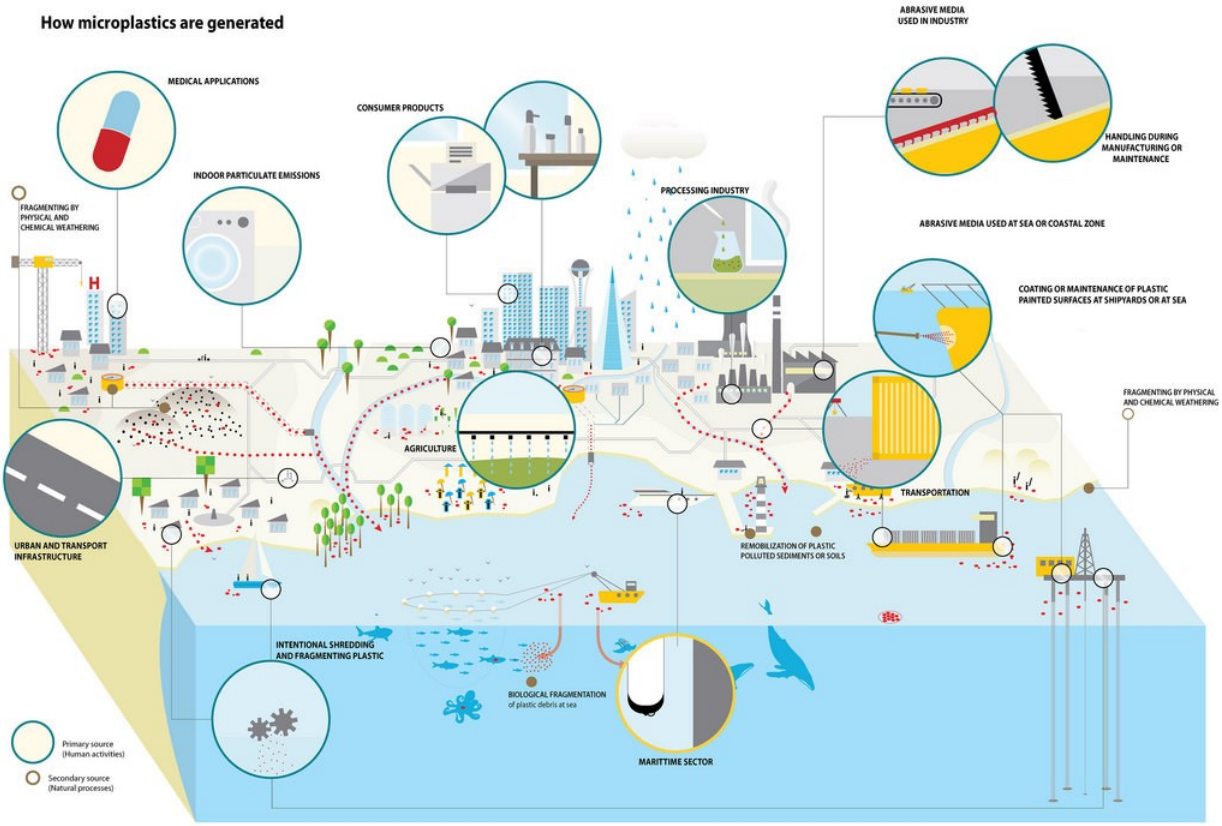


Figure 5. Overview of the major sources of primary microplastics and the generation of secondary microplastics (Source: http://www.grida.no/resources/6929; 2019; GRID-Arendal & Maphoto/Riccardo Pravettoni).

Primary microplastics in terrestrial environments mostly originate from the use of sewage sludge containing microplastics from personal care or household products. In agricultural soils the application of sewage sludge from municipal wastewater treatment plants to farmland is probably a major input, based on recent MP emission estimates in industrialized countries. Plastic pollution in terrestrial systems is also linked to the use of agricultural plastics, such as polytunnels and plastic mulches. Secondary microplastics originate from varying and diverse sources, for example from mismanaged waste either accidentally or intentionally.

- *Effects:*

As plastics have become widespread and ubiquitous in the environment, they are present in a diversity of habitats and can impact organisms at different levels of biological organization, possibly leading to population, community and ecosystem effects. Entanglement is one of the most obvious and dramatic physical impacts of **macroplastics**, as it often leads to acute and chronic injury or death. In particular the higher taxa (mammals, reptiles, birds and fish) are affected, and it may be critical for the success of several endangered species. Because of similar size characteristics to food, plastics are both intentionally and unintentionally ingested by a wide range of species, such as invertebrates, fish, birds and mammals. Ingestion of the non-nutritional plastics can cause damage and/or obstruction of the digestive tract and may lead to decreased foraging due to false feelings of satiation, resulting in reduced energy reserves.

Microplastics and in particular **nanoparticles** that are small enough to be taken up and translocated into tissues, cells and body fluids can cause cellular toxicity and pathological changes due to particle toxicity. In addition, there are also chemical risks involved as plastics can be a source of hazardous chemicals. These chemicals can be part of the plastic itself (i.e. monomers and additives) and/or chemicals that are absorbed from the environment into the matrix or such as lead, cadmium, mercury, persistent organic pollutants (POPs) like PAHs, PCBs and dioxins. However, as this process depends on the fugacity gradients, there is a lot of uncertainty about the extent that transfer of pollutants does occur in the environment. Actually, when taking all exposure pathways into account, the transfer from (micro)plastics seems to be a minor pathway.

Watch the video on the research of Inneke Hantoro.

Finally, marine plastics may act as floating habitats for invasive species, including harmful algal blooms and pathogens, leading to spreading beyond their natural dispersal range and creating the risk of disrupting ecosystems of sensitive habitats.

2.2.9. Nanomaterials

Author: Martina Vijver

Reviewers: Kees van Gestel, Frank van Belleghem, Melanie Kah

Leaning objectives:

You should be able to:

- explain the differences between nanomaterials and soluble chemicals.
- describe nano-specific features and explain the difference between nanomaterials and particles with a larger size, as well as how they differ from molecules.

Keywords: Nanomaterials , emerging technologies, colloids, nanoscale, surface reactivity

Introduction

Engineering at the nanoscale (i.e. 10^{-9} m) brings the promise of radical technological development. Due to their unique properties, engineered nanomaterials (ENMs) have gained interest from industry and entered the global market. Potentials ascribed to nanotechnology are amongst all: stronger materials, more efficient carriers of energy, cleaner and more compact materials that allow for small yet complex products. Currently, nanomaterials are used in numerous products, although exact numbers are lacking. In 2014, the market was estimated to contain more than 13,000 nano-based products (Garner and Keller, 2014). There is a wide variety of products containing nanomaterials, ranging from sunscreens and paint, to textiles, medicines, electronics covering many sectors (Figure 1).

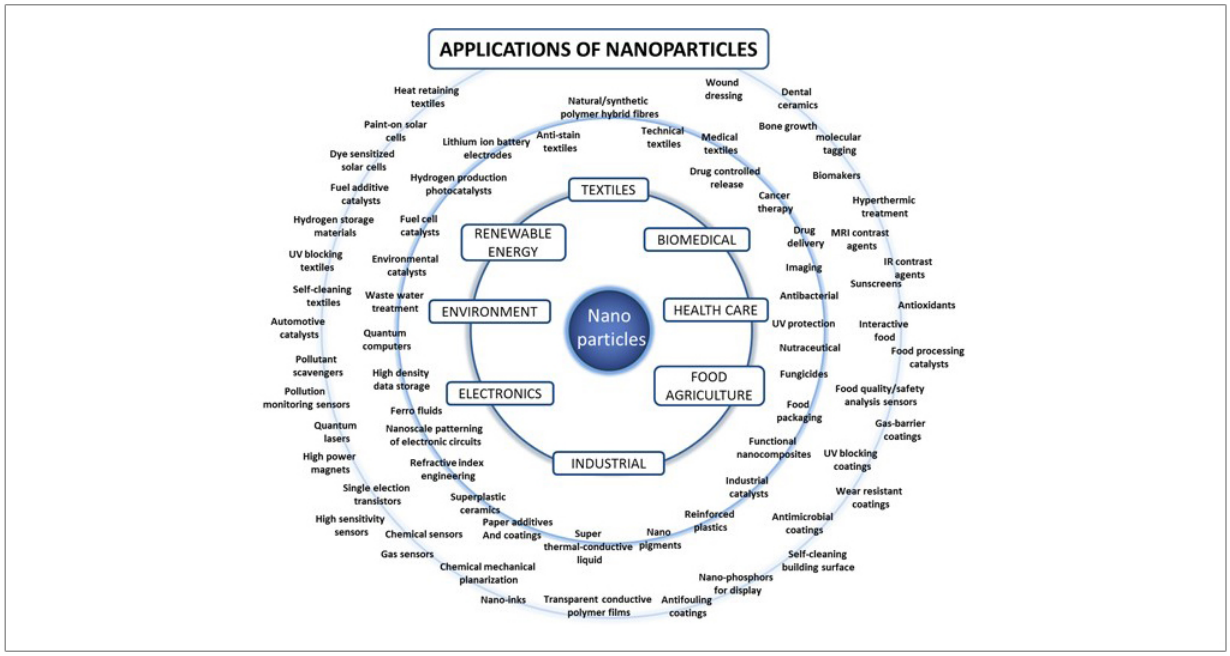


Figure 1. Applications in different sectors where engineered nanomaterials are used (source: www.enteknomaterials.com/wp-c...lzemeler-5.jpg).

Nanomaterials:

The European Commission in 2011 adopted a new definition of *'nanomaterial'* reading *'a natural, incidental or manufactured material containing particles, in an unbound state or as an agglomerate or as an aggregate and where, for 50 % or more of the particles in the number size distribution, one or more external dimensions is in the size range 1 nm-100 nm'.*

Nanomaterials occur naturally, think of minuscule small fine dust, colloids in the water column, volcanic ash, carbon black and colloids known as ocean spray. In paints the features of the colloids are used to obtain the pigment colors. From the year 2000 on, an exponential growth was seen in their synthesis due to the advanced technologies and imaging techniques needed to work on a nano-scale. **First generation nanotechnologies** (before 2005) generally refers to nanotechnology already on the market, either as individual nanomaterials, or as nanoparticles incorporated into other materials, such as films or composites. Surface engineering has opened the doors to the development of second and third generation ENMs. **Second generation nanotechnologies** (2005-2010) are characterised by nanoscale elements that serve as the functional structure, such as electronics featuring individual nanowires. From 2010 onward there has been more research and development of **third generation nanotechnologies**, which are characterised by their multi-scale architecture (i.e. involving macro-, meso-, micro- and nano-scales together) and three-dimensionality, for applications like biosensors or drug-delivery technologies modelled on biological templates. Self-assembling bottom-up techniques have been widely developed at industrial scale, to create, manipulate and integrate nanophases into more complex nanomaterials with new or improved technological features. Post 2015, the **fourth generation** ENMs are anticipated to utilise 'molecular manufacturing': achieving multi-functionality and control of function at a molecular level. **Nowadays, virtually any material can be made on the nanoscale**.

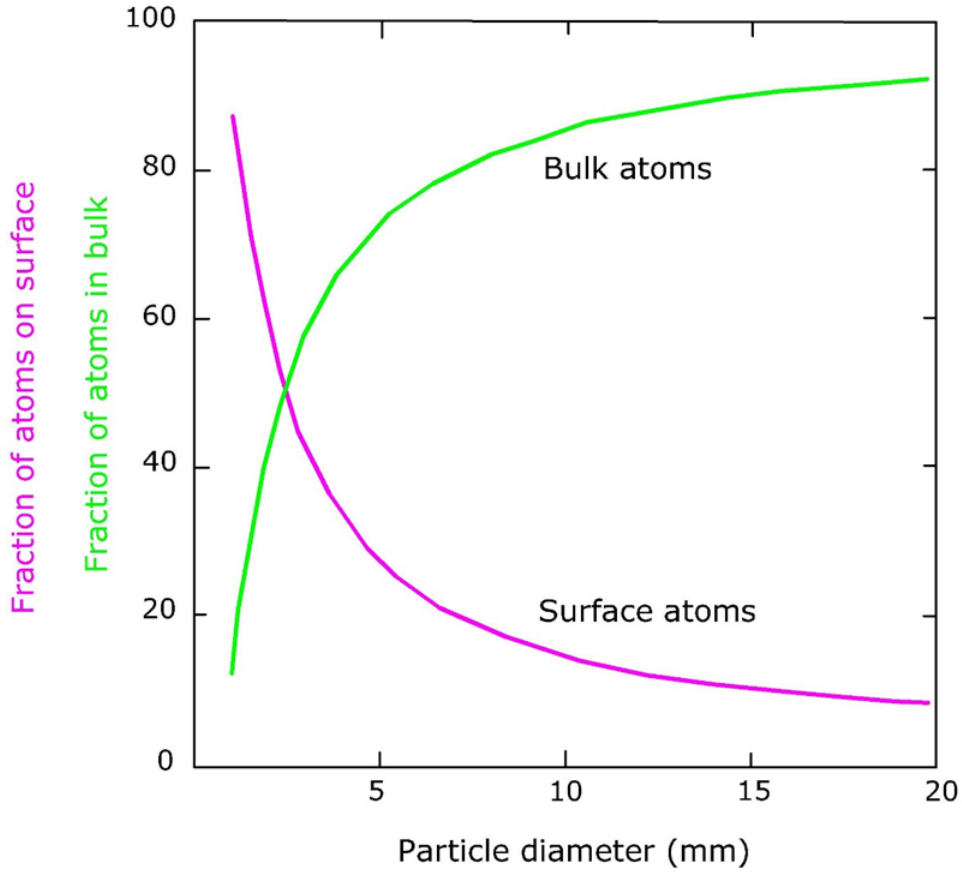


Figure 2. Relationship between particle diameter and the fraction of atoms at the surface. Drawn by Wilma Ijzerman.

Size does matter

Nanoscale materials have far larger surface areas than larger objects with similar masses. A simple thought experiment shows why nanoparticles have phenomenally high surface areas.

A solid cube of a material 1 cm on a side has 6 cm^2 of surface area, about equal to one side of half a stick of gum. When the 1 cm^3 is filled with micrometer-sized cubes - a trillion (10^{12}) of them, each with a surface area of 6 square micrometers - the total surface area amounts to 6 $m^{2.}$. As surface area per mass of a material increases, a greater proportion of the material comes into contact with surrounding materials (Figure 2). Small particles also give that there is a high proportion of surface atoms, high surface energy, spatial confinement and reduced imperfections (Figure 2). It results in the fact that ENMs are having an enlarged reactivity.

compared to larger "bulk" materials. For instance, ENMs have the potency to transfer concentrated medication across the cell membranes of targeted tissues. By engineering nanomaterials, these properties can be harnessed to make valuable new products or processes.

ENMs are often designed to accomplish a particular purpose, taking advantage of the fact that materials at the nanoscale have different properties than their larger-scale counterparts.

ENMs and environmental processes

ENMs are described as a population of particles, and quantified by the particle size distribution (PSD). Nonetheless often a single value (e.g. average ± standard deviation) is reported and not the full PSD. When the particles are suspended in an exposure medium, the size distribution of the NPs is changing over time. After being emitted into aquatic environments, NPs are subject to a series of environmental processes. These processes include dissolution and aggregation (see Figure 3) and subsequent sedimentation. It is known that the behavior and fate of NPs are highly dependent on the water chemistry. In particular, environmental parameters like pH, concentration and type of salts (especially divalent cations) and natural organic matter (NOM) can strongly influence the behaviour of NPs in the environment. For example, pH can affect the aggregation and dissolution of metallic NPs by influencing the surface potential of the NPs (von der Kammer et al., 2010). The divalent cations Ca^{2+} and Mg^{2+} are able to efficiently compress the electrical double-layer of NPs and consequently enhance homo-aggregation and hetero-aggregation of NP (see Figure 2) and the cations will be related to bridging the electrostatic interactions. In surface water, aggregation processes most often lead to sedimentation and sometimes to floating aggregates (depending on the density).

Coating of ENMs will change the dynamics of these processes. As a result of these nano-specific features, ENMs form a suspension which is different from chemicals that dissolve and form a solution. These ENMs in suspension then follow different environmental fate and behaviour compared to solutions. For this reason the way the dosage of ENPs should be expressed is highly debated within the nano-safety community. Should this be on a mass basis, as is the case for molecules of conventional chemicals (e.g. mg/L), or is the particle number the preferred dose metric such as in colloidal science (e.g. number of particles/L, or relative surface-volume ratio) or multi-mixed dosimetry expression. How to express the dose for nanomaterials is a quest still debated within the scientific community (Verschoor et al., 2019).

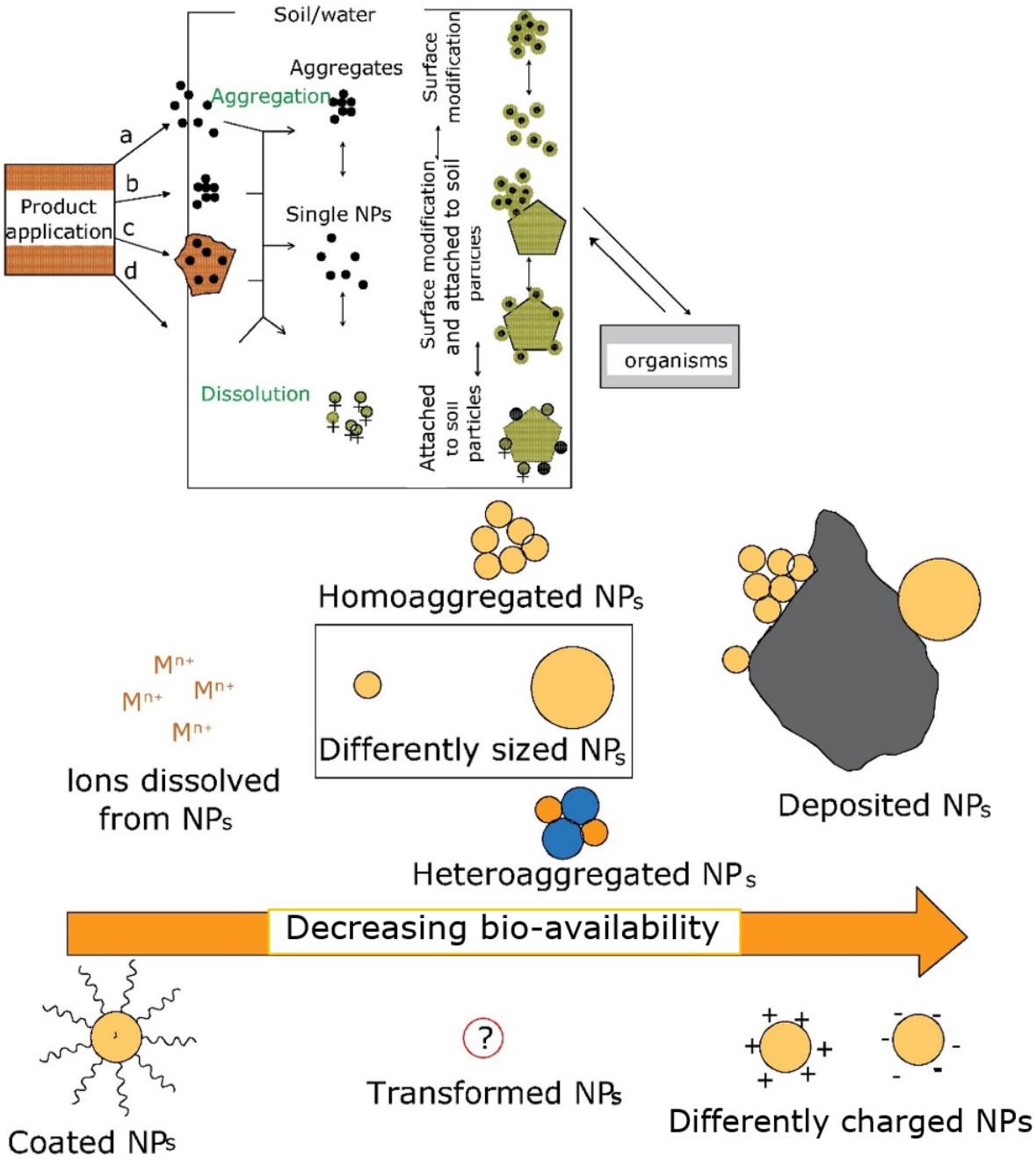


Figure 3.

Top: Schematic illustration how metallic nanoparticles (NPs) behave in an exposure matrix (redrawn from Nowak & Bucheli, 2007). Bottom: Aggregation is divided into two type: homo-aggregation (i.e., aggregation between nanoparticles) and hetero-aggregation (i.e., aggregation of nanoparticle and biomass). Drawn by Wilma IJzerman.

Classification of NMs

Although we learned from the text above that changing the form of a nanomaterial can produce a material with new properties (i.e. a new nanomaterial); often a group of materials developed is named after the main chemical component of the ENMs (e.g. nanoTiO_2) that is available in different (nano)forms. Approaches to group ENMs have been presented below:

- *Classification by dimensionality / shape / morphology:*

Shape-based classification is related to defining nanomaterials, and has been synopsized in the ISO terminology.

- *Classification by composition / chemistry:*

This approach groups nanomaterials based on their chemical properties.

- *Classification by complexity / functionality:*

The nanomaterials that are in routine use in products currently are likely to be displaced by nanomaterials designed to have multiple functionalities, so called 2nd-4th generation nano-materials.

- *Classification by biointerface:*

A proposal relates to the hypothesis that nanomaterials acquire a biological identity upon contact with biofluids and living entities. Systems biology approaches will help identify the key impacts and nanoparticle interaction networks.

References

Garner, K.L., Keller, A.A. (2014). Emerging patterns for engineered nanomaterials in the environment: a review of fate and toxicity studies. Journal of Nanoparticle Research 16: 2503.

Nowack, B., Bucheli, T.D. (2007). Occurrence, behavior and effects of nanoparticles in the environment. Environmental Pollution 150, 5-22.

Verschoor, A.J., Harper, S., Delmaar, C.J.E., Park, M.V.D.Z., Sips, A.J.A.M., Vijver, M.G., Peijnenburg, W.J.G.M. (2019). Systematic selection of a dose metric for metal-based nanoparticles. NanoImpact 13, 70-75.

Von der Kammer, F., Ottofuelling, S., Hofmann, T. (2010). Assessment of the physico-chemical behavior of titanium dioxide nanoparticles in aquatic environments using multi-dimensional parameter testing. Environmental Pollution 158, 3472-3481.

2.2.9. Question 1

What gave us (as mankind) the capacity to virtually synthesize every material at the nanoscale?

2.2.9. Question 2

Explain the statement 'size does matter' for reactivity, use within your explanation the relationship between particle size and the surface area.

2.2.9. Question 3

Give an example of a nano-specific property - which is not size - that enhances reactivity of a material.

This page titled 2.3.2.2: Pollutants with specific properties is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **2.2: Pollutants with specific properties** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.2.3: Pollutants with specific use

2.3. Pollutants with specific use

2.3.1. Crop Protection Products

Author: Kees van Gestel

Reviewers: Steven Droge, Peter Dohmen

Leaning objectives:

You should be able to:

- describe the role of crop protection products in agriculture
- mention different types of pesticides and their different target groups.
- distinguish and mention important chemical groups of pesticides
  - related to the chemistry
  - related to the mode of action
- describe major components included in a commercial formulation of a crop protection product beside the active substance(s).

Keywords: Insecticides, Herbicides, Fungicides, Active substances, Formulations

Introduction

Crop protection products are used in agriculture. The principle target of agriculture is the provision of food. For this purpose, agriculture aims to reduce the competition by other (non-crop) plants and the loss of crop due to herbivores or diseases. An important tool to achieve this is the use of chemicals, such as crop protection products (CPP). Accordingly, CPPs are intentionally introduced into the environment and represent one of the largest sources of xenobiotic chemicals in the environment. These chemicals are by definition effective against the target organism, often already at fairly low doses, but may also be toxic to non-target organisms including humans. The use of pesticides, also named Crop Protection Products (CPP) or often also Plant Protection Products (PPP, the latter term may be misleading for herbicides which are intended to reduce all but the crop plants), is therefore strictly regulated in most countries. The main pesticides used in the largest volumes world-wide are herbicide all s, insecticides, and fungicides. As shown in Table 1, pesticides are used against a large number of diseases and plagues.

Table 1. Classification of pesticides according to what they are supposed to control

Pesticide type	**Target**
acaricides	against mites and spiders (incl. miticides)
algicides	against algae
althelmintics (vermicides)	against parasites
antibiotics	against bacteria and viruses (incl. bactericides)
bactericides	against bacteria
fungicides	against fungi
herbicides	against weeds
insecticides	against insects
miticides	against mites
molluscicides	against slugs and snails
nematicides	against nematodes
plant growth regulators	retard or accelerate the growth of plants
repellents	drive pests (e.g. insects, birds) away
rodenticides	against rodents

Formulations

A pesticidal product usually consists of one or more ***active substances***, that are brought onto the market in a commercial formulation (spray powder, granulate, liquid product etc.). The ***formulation*** is used to facilitate practical handling and application of the chemical, but also to enhance its effect or its safety of use. The active substance may, for instance, be a solid chemical, while application requires it to be sprayed. Or the active substance degrades fast under the influence of sunlight and therefore has to be encapsulated. One of the most used types of formulation is a concentrated emulsion, which may be sprayed directly after dilution with water. In this formulation, the active substance is dissolved in an oily matrix and a detergent is added as emulsifier to make the oil miscible with water. In this way, the active substance becomes quickly available after spraying. In so-called slow-release formulations, the active substance is encapsulated in permeable microcapsules, from which it is slowly released. Another component of a formulation can be a ***synergist***, which increases the efficacy of the active substance, for instance by blocking enzymes that metabolize the active substance. Here an overview of main formulation constituents:

- Solvents: to ease handling and application of the active substances, they are usually dissolved. For a highly water soluble compound this solvent may just be water; however, most compounds have low water solubility and they are thus dissolved in organic solvents.
- Emulsifier, detergents, dispersants: used to provide a homogeneous mixture of the active substance in the aqueous spray solution.
- Carrier: solid formulations, such as wettable granules (WG). All wettable powder (WP) formulations often use inert materials such as clay (kaolinite) as carrier.
- Wetting agent: they help providing a homogeneous film on the plant surface.
- Adjuvant: may help to increase uptake of the active substance into the plant.
- Minor constituents:
  - *Antifreeze agent*: to keep the formulation stable also in cold storage conditions.
  - *Antifoam agent*: some of the formulants may result in foaming during application, which is not wanted.
  - *Preservative, biocide*: to prevent biological degradation of the active substance or the formulants during storage.
  - There are numerous additional specific additional constituents for specific purposes such as *colours*, *detterents*, *stickers*, etc

Four types of nomenclature are used in case of pesticides:

1. The ***trade name***, e.g. Calypso®, which is given by the manufacturer. The same active substance is often sold under more than one different trade names (accordingly, the use of trade names only is not a sufficient description of the test substance in scientific literature).

2. The code name, which is the ***"common" name*** of the active substance. Calypso® 480 SC, for example, is a concentrated suspension containing 480 g/L of the active substance thiacloprid.

3. The ***chemical name*** of the active substance. Thiacloprid is [3-[(6-chloropyridin-3-yl)methyl]-1,3-thiazolidin-2-ylidene]cyanamide.

4. The name of the chemical group to which the active substance belongs, in case of thiacloprid: neonicotinoids.

Chemical classes

Pesticides represent quite a number of different groups of chemicals. Pesticides include inorganic chemicals (like copper used as a fungicide), organic synthetic chemicals, and biologicals (organic natural compounds). Pesticides from the same chemical group may be used against different pest organisms, like the organotin compounds (see below). Some chemicals have a broad mode of action: many soil disinfectants, such as metam-sodium, kill nematodes, fungi, soil insects and weeds. Other pesticides are more selective, like neonicotinoids acting only on insects, or very selective, like the insect-growth regulator fenoxycarb, which is used against leaf-rollers without affecting its natural enemies. Selectivity of a pesticide also indicates to what extent non-target species may be affected upon its application (side-effects). *Integrated pest management* (IPM) aims at an as sustainable as possible crop protection system by combining biological agents (predators of the pest organism) using chemicals having a selective mode of action. Such systems are nowadays receiving increasing interest in different agricultural crops.

Some groups of pesticides that were used or still are widely used are presented in more detail. Their modes of action are discussed in Chapter 4.

- Chlorinated hydrocarbons

Best known representative of this group is DDT (dichloro diphenyl trichloroethane; Figure 1), which was discovered in 1939 by the Swiss entomologist Paul Hermann Müller and seemed to be an ideal pesticide: it was effective, cheap and easy to produce and remained active for a long period of time. As a remedy against Malaria and other insect borne diseases, it has saved millions of human lives. However, the high persistency of DDT, its strong bioaccumulation and its effects on bird populations have triggered the search for alternatives and its ban in most Western countries. But in some developing countries, because of a lack of suitable alternatives for an effective control of malaria, DDT is still in use to kill malaria mosquitos.

Other representatives of chlorinated hydrocarbons are lindane, also called gamma-hexachlorocyclohexane (Figure 1), and the cyclodienes that include the "drins" (aldrin, dieldrin, endrin, See Section 2.1) and endosulfan (Figure 1). Because of their high persistence and bioaccumulative potential, most organochlorinated pesticides have been banned.

Volatile halogenated hydrocarbons were often used as soil disinfectant. These compounds were injected into the soil, and acted as a nematicide but also killed fungi, soil insects and weeds. An example is 1,3-dichloropropene (Figure 1).

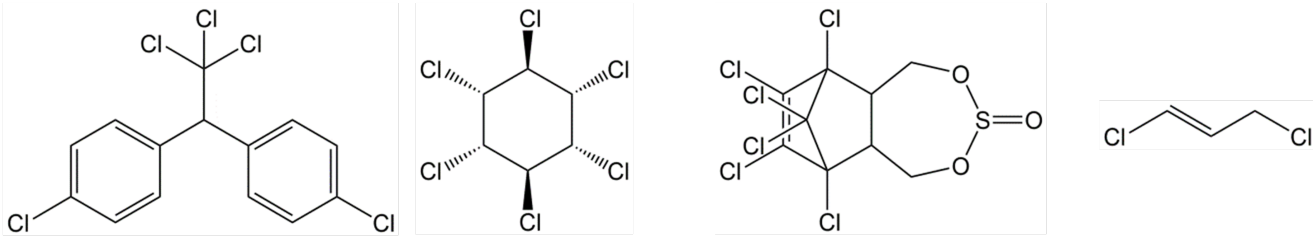


Figure 1. Chemical structures of four different organochlorinated pesticides widely used in the past, from left to right: DDT: 1,1'-(2,2,2-trichloroethane-1,1-diyl)bis(4-chlorobenzene), lindane: gamma-1,2,3,4,5,6-hexachlorocyclohexane, endosulfan:

6,7,8,9,10,10-hexachloro-1,5,5a,6,9,9a-hexahydro- 6,9-methano-2,4,3-benzodioxathiepine-3-oxide, and 1,3-**dichloropropene**. (Source: Steven Droge)

- Organophosphates

Organophosphates are esters of phosphoric acid and constitute important biological molecules such as nucleic acids (DNA) or ATP. Within the contents of pesticides this refers mainly to a group of organophosphate molecules which interfere with acetylcholinesterase. Nerve gases, produced for chemical warfare (e.g., Sarin), also belong to the organophosphates. They are much less persistent and were therefore introduced as alternatives for the chlorinated hydrocarbons. The common molecular structure of organophosphates is a tri-ester of phosphate, phosphonate, phosphorthionate, phosphorthiolate, phosphordithionate or phosphoramidate (Figure 2). With two of the three ester bonds, a methyl- or ethyl- group is bound to the P atom, while the third ester bond binds the rest group or "leaving group".

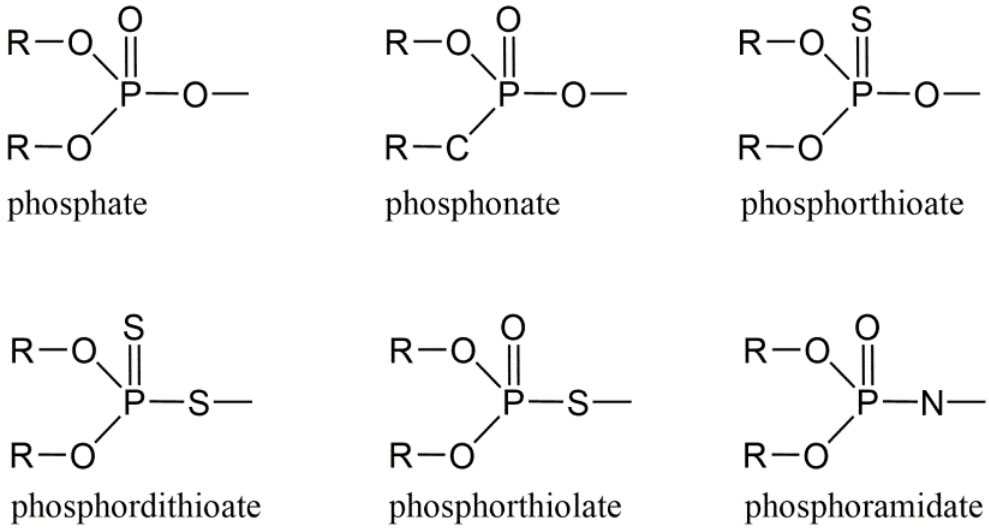


Figure 2: Chemical structure of organophosphates. R = methyl of ethyl group. (Source: Steven Droge)

Dependent on the identity of the latter group, three sub-groups may be distinguished:

1. Aliphatic organophosphates, including malathion (Figure 3) and a number of systemic chemicals.

2. Phenyl-organophosphates, which are more stable than the aliphatic ones but also less soluble in water, like parathion (no longer allowed in Europe; Figure 3).

3. Heterocyclic organophosphates, including chemicals with an aromatic ring containing a nitrogen atom like chlorpyrifos (Figure 3).

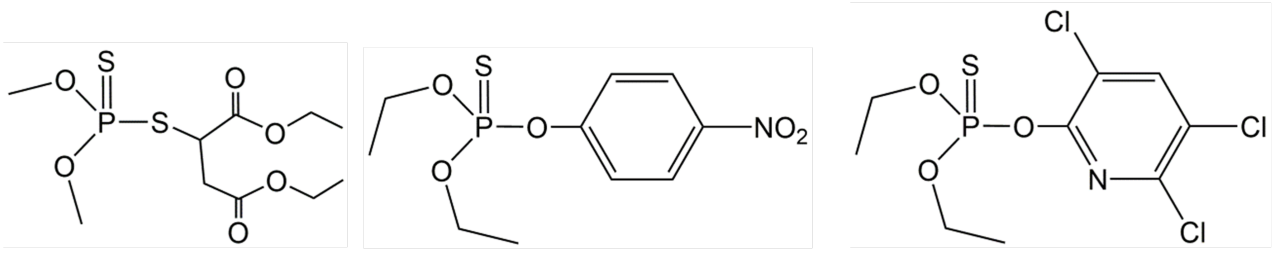


Figure 3: Malathion: diethyl 2-[(dimethoxyphosphorothioyl)sulfanyl]butanedioate (left), parathion: O,O-diethyl-O-4-nitrophenyl-phosphorthioate (middle), and chlorpyrifos: O,O-diethyl O-3,5,6-trichloropyridin-2-yl phosphorothioate (right). (Source: Steven Droge)

- *Carbamates*

Where organophosphates are derived from phosphoric acid, carbamates are derived from carbamate (Figure 4). Their mode of action is similar to that of the organophosphates. The use of older representatives of this group, like aldicarb, carbaryl, carbofuran and propoxur, is no longer allowed in Europe, but diethofencarb (Figure 4), oxamyl and methomyl are still in use.

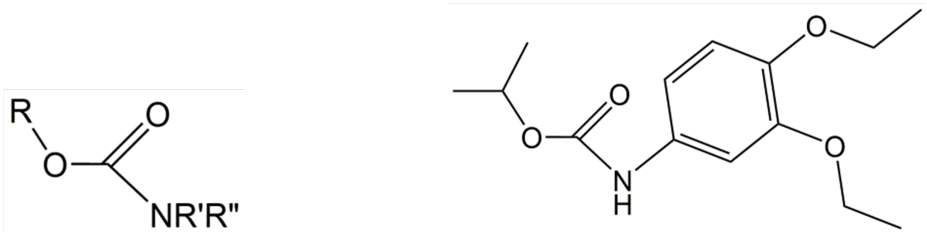


Figure 4: Basic structure of carbamates (left) and diethofencarb: isopropyl 3,4-diethoxycarbanilate (right). (Source: Steven Droge)

- *Pyrethroids*

A number of modern pesticides are derived from natural products. Pyrethroids are based on pyrethrum, a natural insecticide from flowers of the Persian ox-eyed daisy, *Chrysanthemum roseum*. Typical for the molecular structure of pyrethroids is the cyclopropane-carboxyl group (the triangular structure), which is connected with an aromatic group through an ester bond (Figure 5). Pyrethrum is rapidly degraded under the influence of sunlight. Synthetic pyrethroids, which are much more stable and therefore used on a large scale against many different insects, include cypermethrin (Figure 5), deltamethrin, lambda-cyhalothrin, fluvalinate and esfenvalerate.

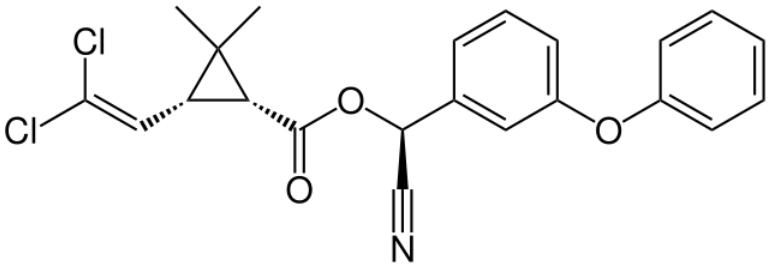


Figure 5: Cypermethrin: [cyano-(3-phenoxyphenyl)methyl]3-(2,2-dichloroethenyl)-2,2-dimethylcyclopropane-1-carboxylate. (Source: Steven Droge)

Neonicotinoids

Based on the natural compound nicotine, which acts as a natural insecticide against plant herbivores, but which was banned as an insecticide due to its high human toxicity, in the 1980s a new group of more specific insecticides has been developed, the neonicotinoids (Figure 6). Several neonicotinoids (e.g., imidacloprid, thiamethoxam) are *systemic*. This means that they are taken up by the plant and exert their effect from inside the plant, either on the pest organism (systemic fungicides or insecticides) or on the plant itself (systemic herbicides). The systemic neonicotinoids are widely applied as seed dressing in major crops like maize and sunflower. Other compounds are mainly used in spray applications, e.g. in fruit growing (thiacloprid, acetamiprid, etc.). Although neonicotinoids are more selective and therefore preferred over the older classes of insecticides like organophosphates, carbamates and pyrethroids, in recent years they have become under debate because of their side effects on honey bees and other pollinators.

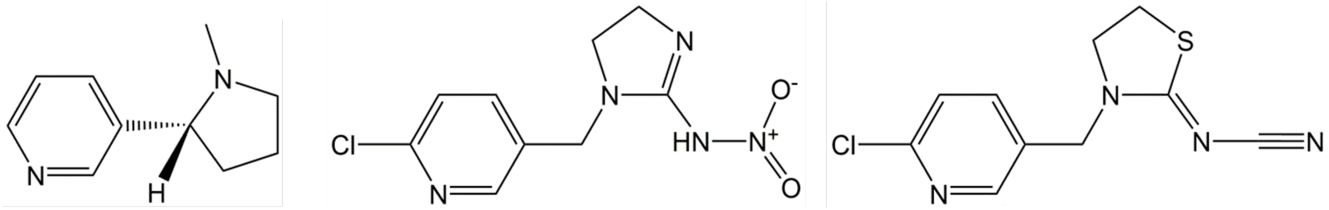


Figure 6: Nicotine: (S)-3-[1-methylpyrrolidin-2-yl]pyridine (left) and the neonicotinoid insecticides imidacloprid: N-{1-[(6-chloro-3-pyridyl)methyl]-4,5-dihydroimidazol-2-yl}nitramide (middle) and thiacloprid: {(2Z)-3-[(6-chloropyridin-3-yl)methyl]-1,3-thiazolidin-2-ylidene}cyanamide (right). (Source: Steven Droge)

- *Isothiocyanates*

Isothiocyanates were used on a large scale as soil disinfectant against nematodes, fungi and weeds. The large number of chemicals with different chemical origin belonging to the isothiocyanates have in common that they form isothiocyanate in soil. A representative of this group is metam-sodium (Figure 7).

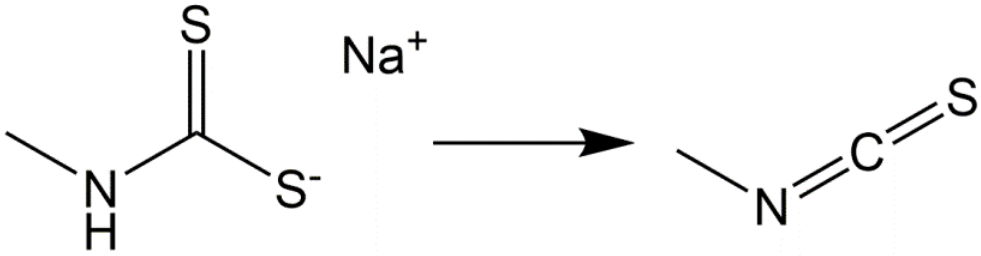


Figure 7. Metam-sodium: sodium methylaminomethanedithioate forming methyl isothiocyanate. (Source: Steven Droge)

- *Organotin compounds*

Fentin hydroxide (Figure 8) was used as a fungicide against *Phytophthora* (causing potato -disease). Tributyltin compounds (TBT) were used as anti-fouling agent (algicide) on ships. TBTC (tributyltin chloride) is extremely toxic to shell-fish, such as oysters, and for this reason banned in many countries. Fenbutatin-oxide was used as an acaricide against spider mites on fruit trees (tributyltin chloride).

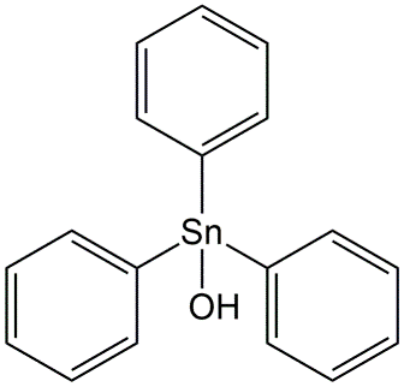


Figure 8: Fentin hydroxide: triphenyltin hydroxide. (Source: Steven Droge)

- *Ryanoids*

Also indicated as diamide insecticides, this group includes chemically distinct synthetic compounds such as chlorantraniliprole (Figure 9), flubendiamide, and cyantraniliprole, that act on the ryanodine receptor and are used against chewing and sucking insects.

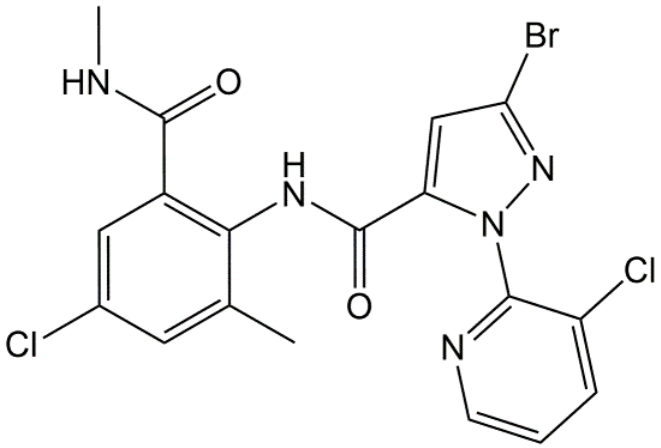


Figure 9: Chlorantraniliprole: 5-bromo-N-[4-chloro-2-methyl-6-(methylcarbamoyl)phenyl]-2-(3-chloropyridin-2-yl)pyrazole-3-carboxamide. (Source: Steven Droge)

- *Phenoxy acetic acids*

Phenoxy acetic acids are systemic herbicides, exerting their action after uptake by the leaf and translocation throughout the plant. Especially plants with broad, horizontally oriented leaves are sensitive for these herbicides. 2,4-D (Figure 10) is the best known representative of this group.

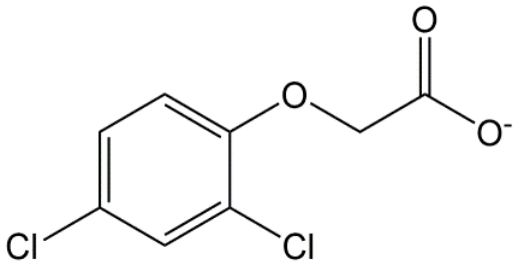


Figure 10. 2,4-D: the anionic form of 2,4-dichloro phenoxy acetic acid (pKa 2.73). (Source: Steven Droge)

- *Triazines*

Triazines are heterocyclic nitrogen compounds, whose structure is characterized by an aromatic ring in which three carbon atoms have been replaced by nitrogen atoms. Triazines are usually applied to the soil before seed germination. The use of several compounds (atrazine, simazine) has been banned in Europe, while others like metribuzin and terbuthylazine (Figure 11) are still in use.

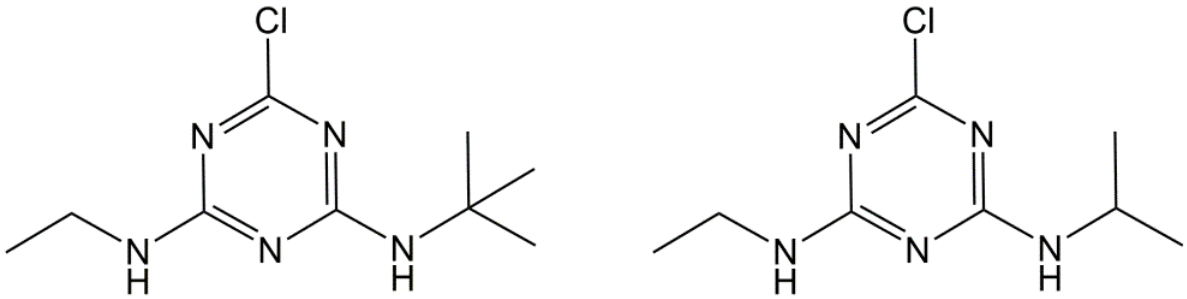


Figure 11. Terbuthylazine: N-tert-butyl-6-chloro-N'-ethyl-1,3,5-triazine-2,4-diamine (left), common replacement of the EU-banned herbicide atrazine (right). (Source: Steven Droge)

- *Bipiridyls*

This group contains the herbicides diquat and paraquat (Figure 12) which mainly act as contact herbicides. This means they damage the plant without being taken up. In soil, they are rapidly inactivated by strong binding to soil particles. The use of paraquat is no longer allowed in Europe, but diquat is still in use.

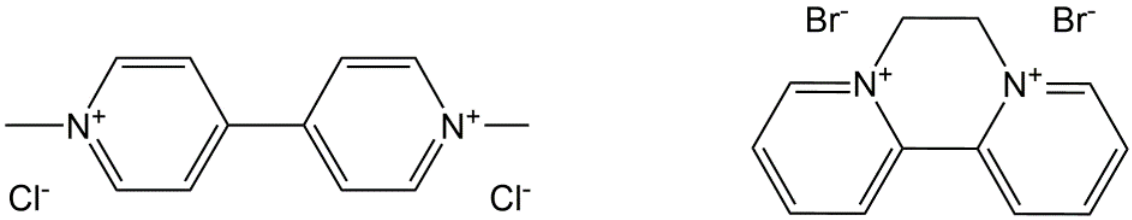


Figure 12. Paraquat: 1,1′-Dimethyl-4,4′-bipyridinium dichloride (left), and diquat: 1,1'-Ethylene-2,2'-bipyridylium dibromide (right). (Source: Steven Droge)

- *Glyphosate and Glufosinate*

As an alternative to the above mentioned herbicides, glyphosate and later glufosinate were developed. These are systemic broad-spectrum herbicides with a relatively simple chemical structure (Figure 13). Their low toxicity to other organisms triggered pesticide producers to introduce *genetically modified crops* (e.g. soybean, maize, oilseed rape, and cotton) that contain incorporated genes for resistance against these broad-spectrum herbicides. This type of resistance allows the farmer to use the herbicide without damaging the crop. For this reason, environmentalist fear an unrestricted use of these herbicides, which indeed is the case especially for glyphosate (better known under the formulation name Roundup®).

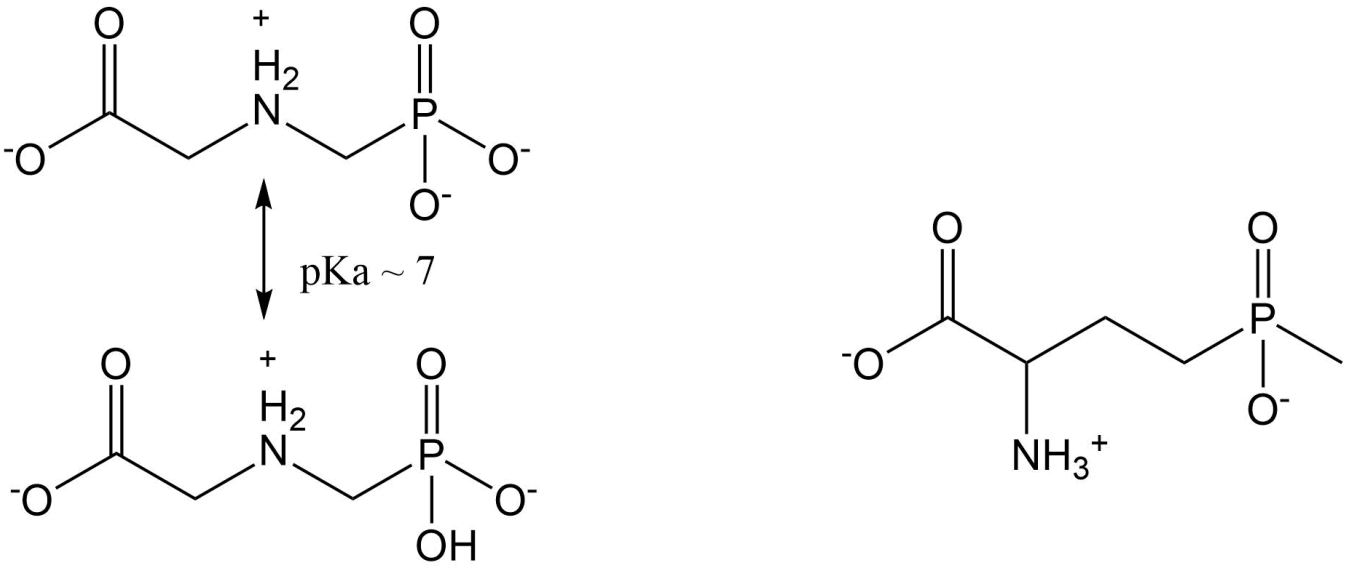


Figure 13. Glyphosate: N-(phosphonomethyl)glycine in the two species most relevant for natural pH range (left), and glufosinate: (RS)-2-Amino-4-(hydroxy(methyl)phosphonoyl)butanoic acid in the most relevant species for natural pH range (right). (Source: Steven Droge)

- Triazoles

Several modern fungicides are sharing a triazole group (Figure 14). These fungicides have gained importance because of problems with the resistance of fungi against other classes of fungicides. Members of this group for instance are epoxiconazole, propiconazole and tebuconazole.

https://bio.libretexts.org/@go/page/108670

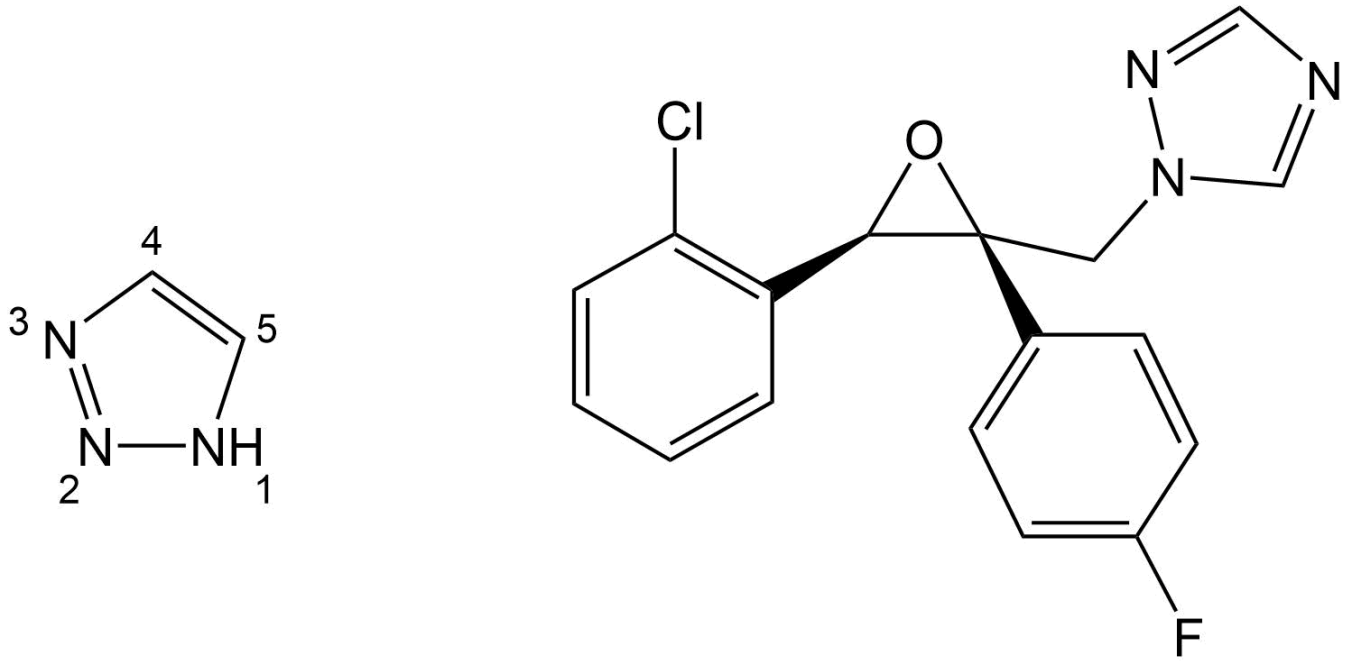


Figure 14: Triazole: 1H-1,2,3-Triazole (left), and epoxiconazole: (2RS,3SR)-1-[3-(2-chlorophenyl)-2,3-epoxy-2-(4-fluorophenyl)propyl]-1H-1,2,4-triazole (right). (Source: Steven Droge)

- *Biological pesticides*

Biological pesticides are produced in living organisms as secondary metabolites to protect themselves against predators, herbivores, parasites or competition. They can be highly effective and act at low concentrations (high toxicity), but in contrast to some synthetic pesticides they are usually sufficiently biodegradable. Compounds like pyrethrum or strobilurin are produced within the plant or within the fungus and are thus protected against photolysis or other environmental degradation. Furthermore, the living organism can produce additional quantities of the secondary metabolite on demand. When used as a pesticide applied as a spray, however, the molecule needs to be modified to enhance its stability (for example against photolysis) to remain sufficiently active over a sufficient period of time. Accordingly, synthetic derivatives of these biological molecules are often more stable, less biodegradable. Examples are the Bt insecticide, which contains an endotoxin highly toxic to insects produced by the bacterium *Bacillus thuringiensis*, and avermectins, complex molecules synthesized by the bacterium *Streptomyces avermitilis*. Avermectins act as insecticides, acaricides and have anthelminthic properties. In nature, eight different forms of avermectin have been found. Ivermectin is a slightly modified structure that is synthesized and marketed commercially. Other compounds belonging to this group are milbemectin and emamectin.

Genetically modified plants containing a gene coding for the toxin produced by the bacterium *Bacillus thuringiensis* (or Bt) are another example of genetic modification being applied in agriculture produce insect-resistant crops.

References:

EU Pesticides Database

2.3.1. Question 1

Systemic pesticides are easily taken up by plants and internally distributed over all plant tissues. What do you expect regarding the water solubility of systemic pesticides?

1. Very Low
2. Very high

2.3.1. Question 2

Why are systemic insecticides applied as seed dressing, so dosed into the soil, dangerous for honey bees and other pollinators?

2.3.1. Question 3

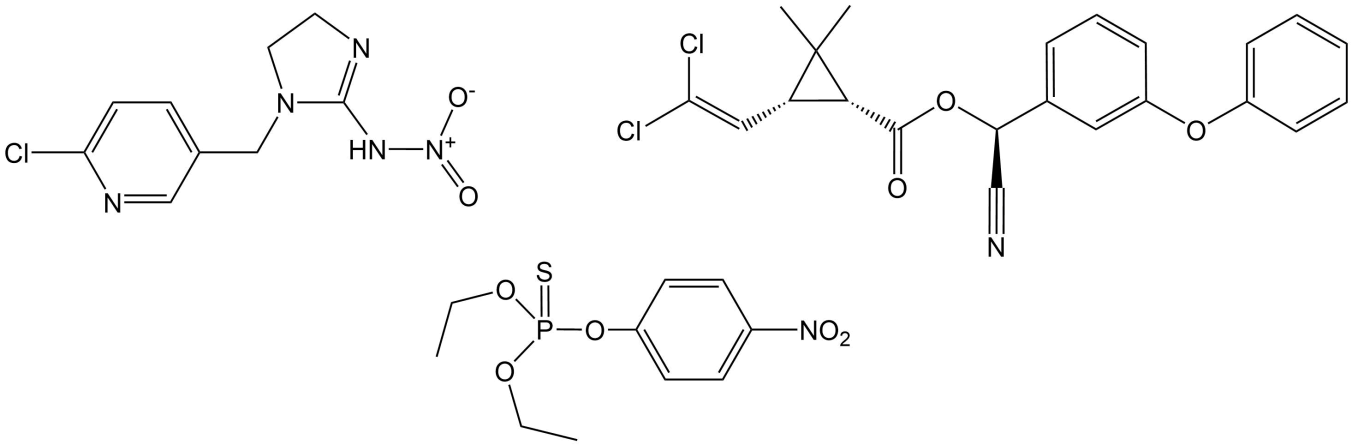
A pesticide with an active substance that is hardly soluble in water is introduced to the market in a commercial formulation. What components should be present in the formulation to allow homogenous application of the pesticide by spraying?

2.3.1. Question 4

Name three important groups of insecticides and mention one typical property or characteristic.

2.3.1. Question 5

1. The figures shows three chemical structures. Indicate which of these chemicals is
    - An organophosphate (circle)
    - A pyrethroid (triangel)
    - A neonicotinoid (square)



Aantal antwoorden: 3

2.3.2. Biocides

Author: Thomas Wagner

Reviewers: Steven Droge, Kevin Thomas

Learning objectives:

You should be able to:

- Understand the purpose of using biocides
- Distinguish different groups of biocides
- Understand the legislation concerning the production and use of biocides
- Understand the potential impact of biocides on ecosystems

Keywords: Biocides, product types, Biocidal Product Regulation (BPR), environmental impact

Introduction

European legislation describes a biocide as '*chemical substance or microorganism intended to destroy, deter, render harmless, or exert a controlling effect on any harmful organism by chemical or biological means*'. The US Environmental Protection Agency (EPA), an independent agency of the U.S. federal government to protect the environment, defines biocides as *'a diverse group of poisonous substances including preservatives, insecticides, disinfectants and pesticides used for the control of organisms that are harmful to human or animal health or that cause damage to natural or manufactured products'*. The definition by the EPA includes pesticides (Chapter 2.3.1). In the scientific and non-scientific literature, the distinction between biocides, pesticides and plant protection products is often vague.

Biocides are used all around us:

- The toothpaste that you used this morning contains biocides to preserve the toothpaste

- The water that you used to flush you mouth is prepared with biocides for disinfection
- The clothes that you are wearing are impregnated with biocides to prevent smells
- The food that you ate for breakfast might have contained biocides to preserve the food
- The construction materials around you have surface coatings that contain biocides to prevent biological degradation of the material

A biocide contains an 'active substance', which is the chemical that is toxic to its target organism, and often contain 'non-active co-substances', which could help in reaching desired product parameters, such as a viscosity, pH, colour, odour or increase its handling or effectiveness. The combination of active substances and non-active substances together makes up the 'biocidal product'. An example of a well-known biocidal product is TriChlor, which contains active substance chlorine that is used to disinfect swimming pools. Because it is impractical to store chlorine gas for the treatment of swimming pools, TriChlor tablets are added to the pool water. TriChlor is trichloroisocyanuric acid (Figure 1). When dissolved in water, the Cl atoms are replaced by H atoms, forming chlorine (Cl^-) and cyanuric acid (Figure 2). The free chlorine is able to disinfect the swimming pool.

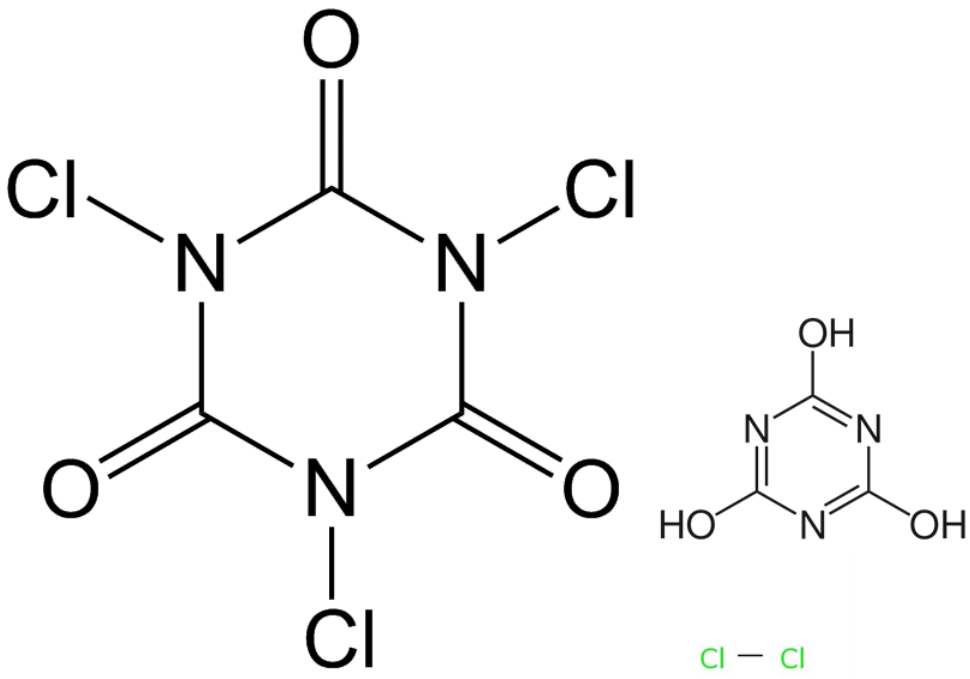
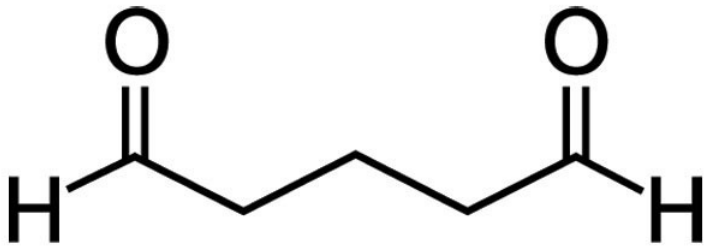


Figure 2. Cyanuric acid and chlorine (B)

A biocidal product can also contain multiple biologically active substances to enhance its effectivity, such as AQUCAR™ 742 produced by DuPont. It contains glutaraldehyde (Figure 3) and quaternary ammonium compounds (Figure 4) that have a synergistic toxic effect on microorganisms that are present in oilfields and could form biofilms in the pipelines.



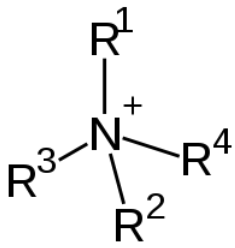


Figure 4. Quaternary ammonium compound

Product types

The biocidal products are classified into 22 different product-types by the European Chemicals Agency (ECHA) (Table 1). It is possible that an active substance can be classified in more than one product types.

Table 1. *The classification of biocides in 22 product types (www.echa.europe.eu)*

Main group 1: Disinfectants and general biocidal products
Product type 1 - Human hygiene biocidal products
Product type 2 - Private area and public health area disinfectants and other biocidal products
Product-type 3 - Veterinary hygiene biocidal products
Product type 4 - Food and feed area disinfectants
Product-type 5 - Drinking water disinfectants
Main group 2: Preservatives
Product-type 6 - In-can preservatives
Product-type 7 - Film preservatives
Product-type 8 - Wood preservatives
Product-type 9 - Fibre, leather, rubber and polymerised materials preservatives
Product-type 10 - Masonry preservatives
Product-type 11 - Preservatives for liquid-cooling and processing systems
Product-type 12 - Slimicides
Product-type 13 - Metalworking-fluid preservatives
Main group 3: Pest control
Product-type 14 - Rodenticides
Product-type 15 - Avidicides
Product-type 16 - Molluscicides
Product-type 17 - Piscicides
Product-type 18 - Insecticides, acaricides and products to control other arthropods
Product-type 19 - Repellents and attractants
Product-type 20 - Control of other vertebrates
Main group 4: Other biocidal products
Product-type 21 - Antifouling products
Product-type 22 - Embalming and taxidermist fluids

Legislation

In Europe, biocides are authorised for production and use by the Biocidal Products Regulation (BPR, Regulation (EU) 528/2012) of the ECHA. The BPR 'aims to improve the functioning of the biocidal product market in the EU, while ensuring a high level of protection for humans and the environment.' (echa.europe.eu/legislation). This is an alternative regulatory framework than that for the plant protection products, managed by the European Food Safety Authority (EFSA). All biocidal products go through an extensive authorisation process before they are allowed on the market. The assessment of a new active starts with the evaluation of a product by the authorities of an ECHA member state, after which the ECHA Biocidal Products Committee forms an opinion. The European Commission then makes a decision to approve or reject the new active substance based on the opinion of ECHA. This approval is granted for a maximum of 10 years and needs to be renewed after it reaches the end of the registration period. The BPR has strict criteria for new active substances, and meeting the following 'exclusion criteria' will result in the new active substance not being approved:

- Carcinogens, mutagens and reprotoxic substances categories 1A or 1B according to CLP regulation

- Endocrine disruptors
- Persistent, bioaccumulative and toxic (PBT) substances
- Very persistent and very bioaccumulative (vPvB) substances

In very special cases, new active substances will be allowed on the market when meeting this exclusion criteria, if they are important for public health and public interest and there are no alternatives available. To lower the pressure on public health and the environment, there is also a candidate list for active substances to be substituted for less harmful active substances when the old active substances meet the following criteria:

- It meets one of the exclusion criteria
- It is classified as a respiratory sensitizer
- Its toxicological reference values are significantly lower than those of the majority of approved active substances for the same product-type and use
- It meets two of the criteria to be considered as PBT
- It causes concern for human or animal health and for the environment even with very restrictive risk management measures
- It contains a significant proportion of non-active isomers or impurities

The impact of environmental release

The release of biocides in the environment can have huge consequences, since these products are designed to cause damage to living organisms. A classic example is the release of tributyltin from shipyards, harbours and on sailing routes from the antifouling paint on the hulls of ships (De Mora, 1996). Tributyltin was used in the antifouling paint from the 1950s on to prevent microorganisms from settling on the hulls of ships, which would increase the fuel costs and repair costs. However, the release of tributyltin from the paint resulted in a toxic effect on organisms at the bottom of the food chain, such as algae and invertebrates. Tributyltin then biomagnified in the food web, this way affecting larger predators, such as dolphins and sea otters. Eventually, tributyltin entered the diet of humans. The first legislation on the use of tributyltin for ships dates back to the 1980s, but it was not until the Rotterdam Convention of 2008 that the complete use of tributyltin as an active biocide in antifouling paints was banned. Biocides can also have an effect on the capability of the environment to deal with pollution. Microorganisms are responsible for cleaning polluted areas by using the pollutant as food-source. McLaughlin et al. (2016) studied the effect of the release of biocide glutaraldehyde in spilled water from hydraulic fracturing on the microbial activity and found that the microbial activity was hampered by the biocide glutaraldehyde. Hence, because of the biocide, the environment was not or slower capable to return to its original state.

References

De Mora, S.J. (1996). Tributyltin: case study of an environmental contaminant, Vol. 8, Cambridge Univ. Press

McLaughlin, M.C., Borch, T., Blotevogel, J. (2016). Spills of hydraulic fracturing chemicals on agricultural topsoil: Biodegradation, sorption and co-contaminant interactions, Environmental Science & Technology 50, 6071-6078

2.3.2. Question 1

Explain the difference between an active substance and a biocidal product.

2.3.2. Question 2

What is the purpose of using non-active co-substances in a biocidal product?

2.3.2. Question 3

A ship is transporting bananas from Costa Rica to the Netherlands. Mention 5 biocide product types that could be used on this ship, and their purpose.

2.3.2. Question 4

Mention 3 potential ways in which the environment can be damaged as the result of the (accidental) release of biocides in common daily practice.

2.3.3. Pharmaceuticals and Veterinary Pharmaceuticals

(draft)

Author: Thomas ter Laak

Reviewer: John Parsons, Steven Droge

Leaning objectives:

You should be able to:

Keywords: emission, immission, waste water treatment, disease treatment

Introduction

Pharmaceuticals are consumed by humans (human pharmaceuticals) and administered to animals (veterinary pharmaceuticals).

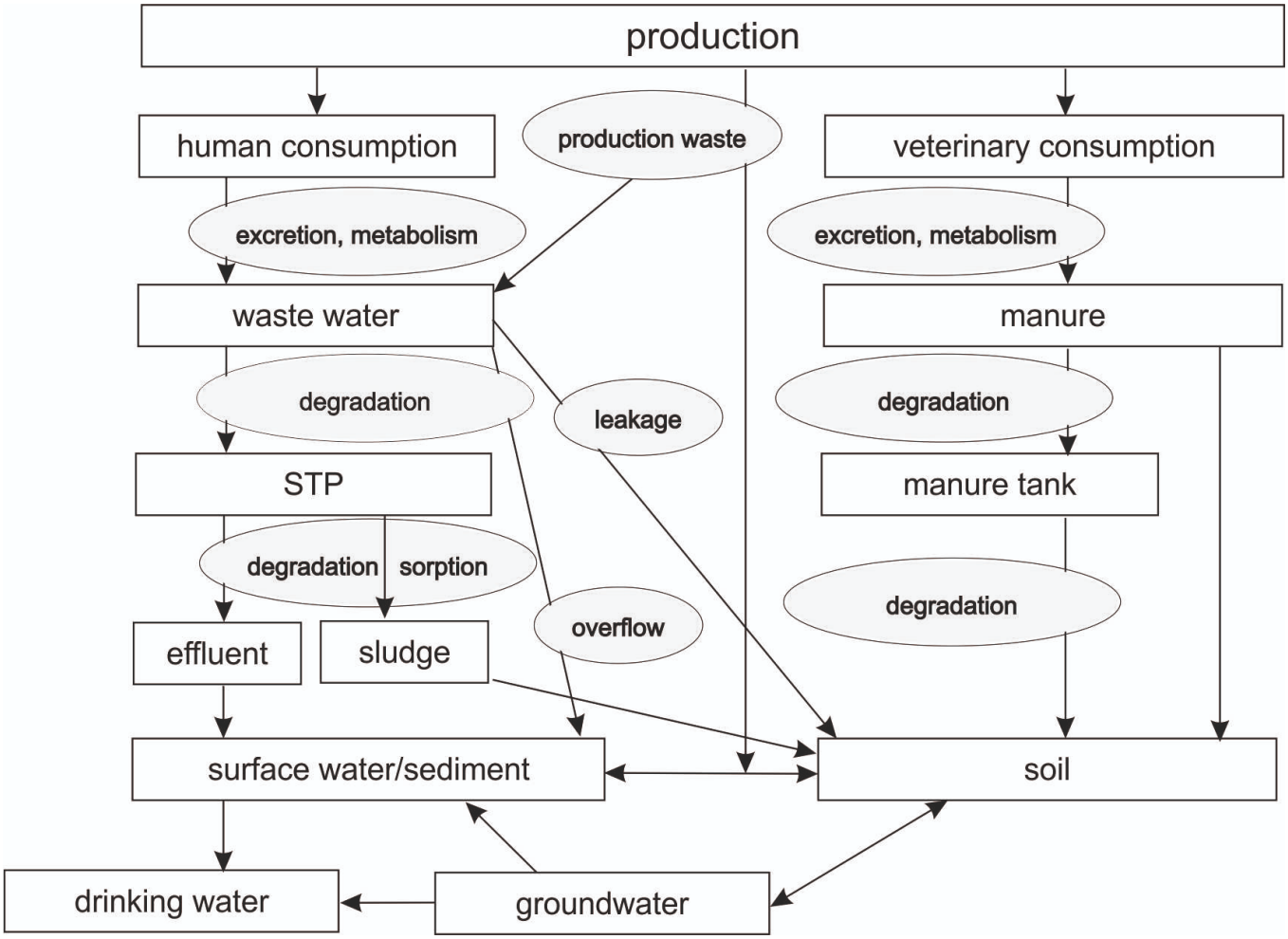


Figure 1. Pharmaceutical emissions routes to the (aqueous) environment. STP = sewage treatment plant. (Source: Steven Droge).

The active ingredients used in human and veterinary medicine partially overlap, however, the major fraction of pharmaceutically active substances in use are restricted to human consumption. In veterinary practice most of the applied pharmaceuticals are antibiotics and anti-parasitic agents, while in human medicine, pharmaceuticals to treat e.g. diabetes, pain, cardiovascular diseases, autoimmune disorders and neurological disorders make up a much larger portion of the pharmaceuticals in use. Worldwide pharmaceutical consumption has increased over the last century (several staggering numbers are summarized here). It is expected that the consumption will further increase due to a wider access to pharmaceuticals in developing countries. Additionally, demographic trends such as aging populations that is often seen in developed countries can also lead to increased consumption of pharmaceuticals, since older generations generally consume more pharmaceuticals than younger generations (van der Aa et al.,

2011). The widespread and increasing use and their biological activity makes them relevant for environmental research. Below an overview is given on the emission, occurrence and fate(modeling) of pharmaceuticals in the environment.

Pharmaceuticals in the environment

Pharmaceuticals can enter the environment through various routes. Figure 1 gives and overview of emission routes of pharmaceuticals to the environment. Pharmaceutical are produced, transported to users (humans and/or animals), partially excreted by the users via urine and feces. For humans, the excrements are transport to wastewater treatment plants, septic tanks or directly emitted to soil or surface water. For animals, and especially livestock, manure contains residues of the pharmaceuticals. These pharmaceuticals end up in the environment when animals are grazing outside or when centrally collected manure is applied as fertilizer on arable land. Just like the pharmaceutical consumption, metabolism in the user (a human or an animal), the treatment and further application of communal wastewater and manure varies. Subsequently, emissions also vary between pharmaceuticals, countries and locations.

Concentrations of pharmaceuticals in the environment strongly vary, in Figure 2 concentration ranges of pharmaceuticals and some of their transformation products in the Meuse river and some tributaries are shown.

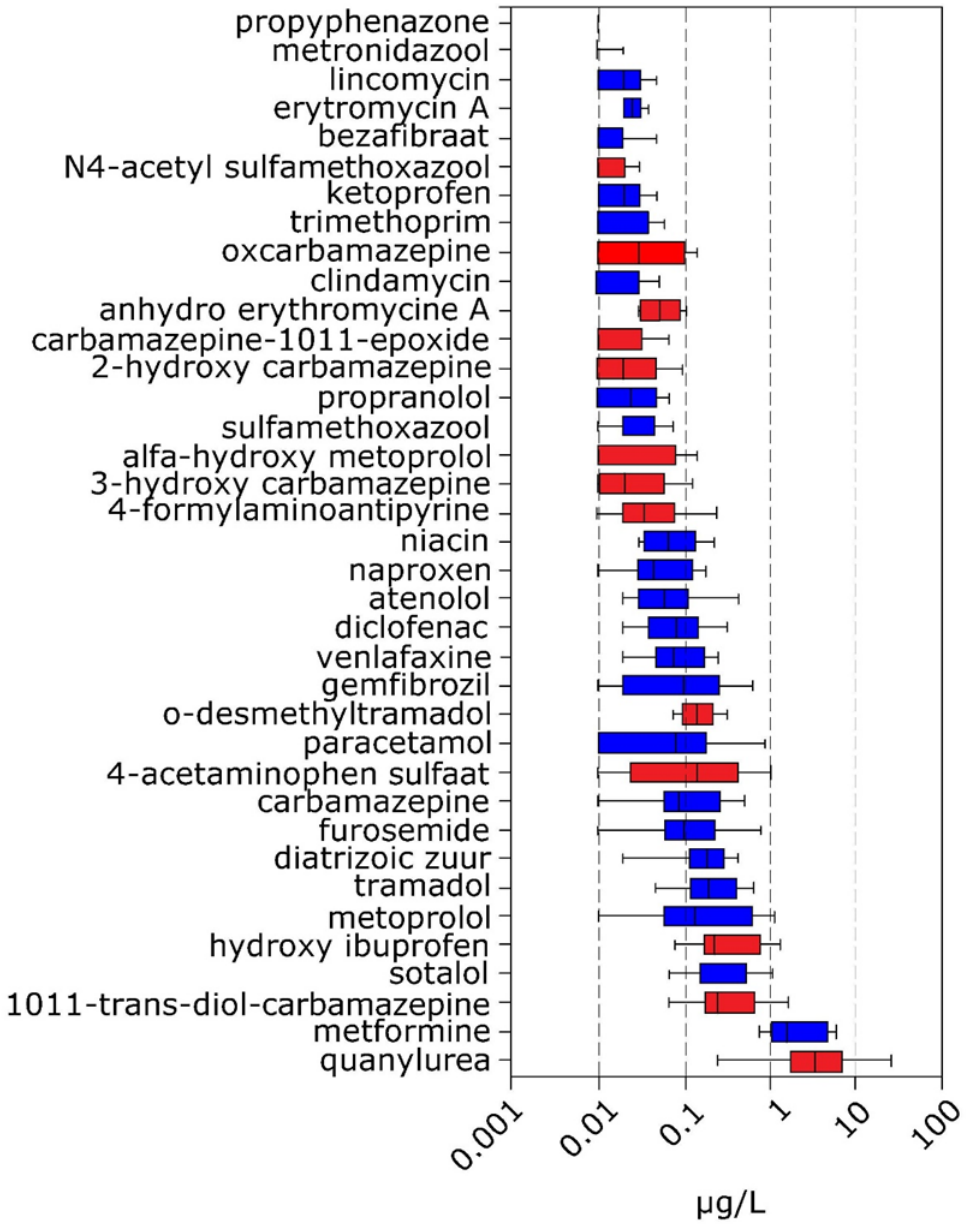


Figure 2. Pharmaceutical concentrations of pharmaceuticals in the River Meuse and some of its tributaries (adapted from Ter Laak et al. 2014). Parent

pharmaceuticals are plotted in bleu, transformation products in red. Drawn by Wilma IJzerman.

Properties of pharmaceuticals and their behavior and fate in the environment

Pharmaceuticals in use are developed for a wide array of diseases and therapeutic treatments. The chemical structures of these substances are therefore also very diverse, considering their size, structural presence of specific atoms, and physicochemical properties such as their hydrophobicity, aqueous solubility and ionization under environmentally relevant pH values, as shown for some examples in Figure 3.

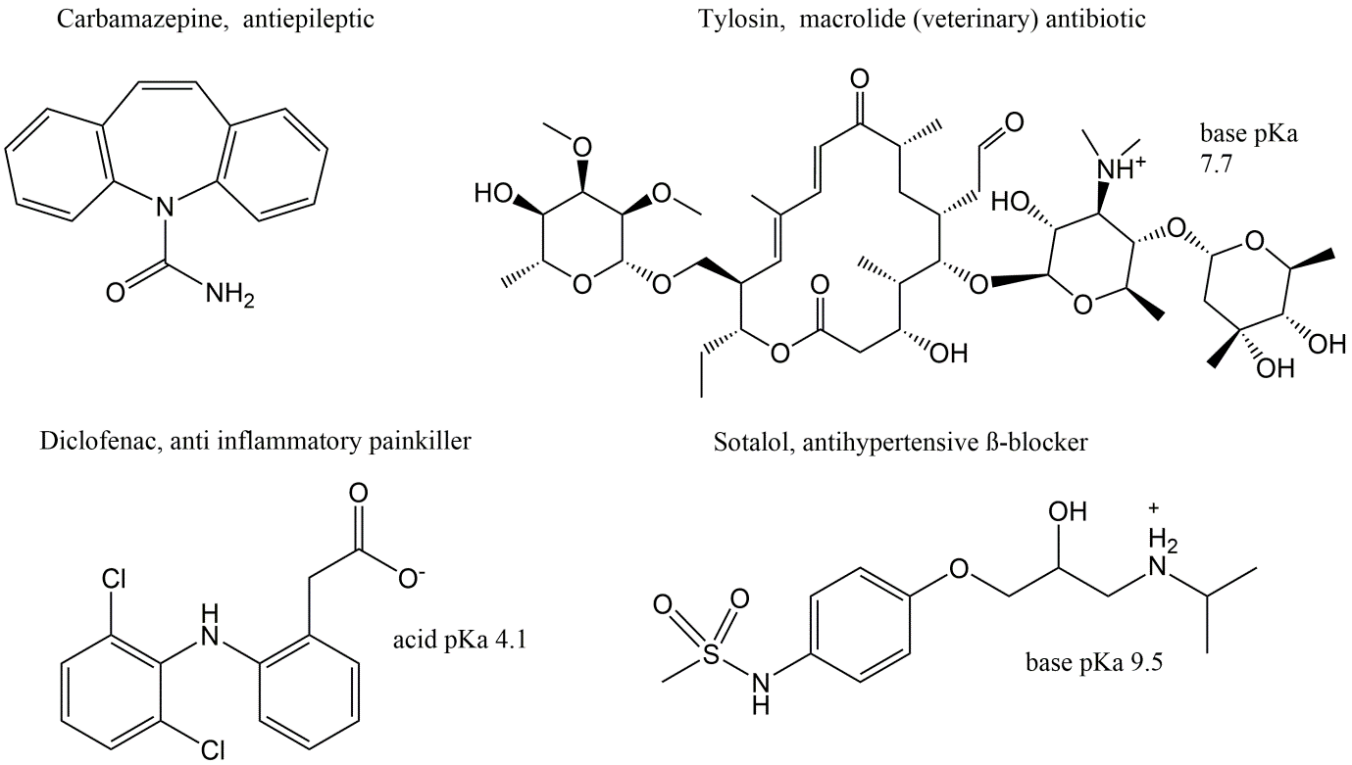


Figure 3. Examples of pharmaceuticals, illustrating the variable chemical structures. (Source: Steven Droge)

As a consequence of their structural diversity their environmental distribution and fate is also very variable. This makes it difficult to generalize the environmental behavior and fate of pharmaceuticals. Nevertheless, pharmaceuticals have generally certain properties in common:

These three generic properties also make them of environmental relevance since:

Occurrence and modelling of human and veterinary pharmaceuticals in the environment

Pharmaceuticals have been studied in the environment since the 1990s. Most studies have been performed in surface waters, but wastewater (effluents), groundwater, drinking water, manure, soil and sediments were also studied. Pharmaceuticals have been observed in all these matrices in concentrations generally varying from µg/L to sub ng/L levels (Aus der Beek et al., 2016, Monteiro and Boxall, 2010). Various studies have related environmental loads and related concentrations to human consumption data. Basically such -so called- mass balance or immision-emission balancing studies work according the following principle:

Modelling pharmaceuticals in the environment

Since the application of human pharmaceuticals, the consumption, is relatively well documented. Environmental concentrations can be related to consumption. This prediction works the best for the most persistent pharmaceuticals, since variations in loss factors are marginal for these pharmaceuticals. When loss factors become larger, they generally also become more variable, trough seasonal variations in use as well as variation in loss during wastewater treatment and loss processes in the receiving rivers, respectively. This makes the loads and concentrations of more degradable pharmaceuticals are more difficult to predict (Ter Laak et al., 2010).

Loads in a particular riverine system (such as a tributary of the river Meuse in the example below) can be predicted with a very simplified model. Here the pharmaceutical consumption over a selected period is multiplied by the factor of the selected pharmaceuticals that are excreted unchanged by the human body (ranging from 0 to 1)and the fraction that is able to pass the wastewater treatment plant (WWTP) (ranging from 0 to 1):

$$Load\ River\ (\frac{kg}{day})\ Human\ Consumption\ (\frac{kg}{day})\ *\ fraction\ Excreted\ *\ fraction\ Passing\ WWTP$$

When this is related to actual measured concentrations and loads calculated from these numbers, the correlation between predicted and measured loads can be plotted. Various studies have shown that environmental loads can be predicted within a factor of 3 for most commonly observed pharmaceuticals (Ter Laak et al., 2014; Ter Laak et al., 2010; Alder et al., 2010; Oosterhuis et al., 2013).

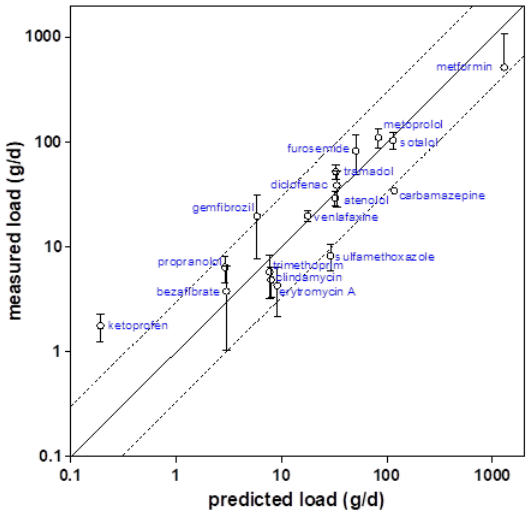


Figure 4. Measured versus predicted loads in a tributary of the Meuse river (adapted from Ter Laak et al., 2014)

For veterinary pharmaceuticals this so called 'immision-emission balancing' is more difficult due to a number of reasons:

In a way the emissions and fate of veterinary pharmaceuticals is similar to emissions of pesticides used in agriculture, but then with a much poorer understanding on loads entering the system and the fate related to the various emission routes and emissions in combination with a complex matrix (urine, feces manure) (Guo et al., 2016). As a consequence, environmental fate studies of veterinary pharmaceuticals often describe specific cases, or cover laboratory studies to unravel specific aspects of the environmental fate of these pharmaceuticals (Kaczala and Blum, 2016; Kummerer, 2009).

Concluding remarks

Pharmaceuticals are commonly found in the environmental compartments such as surface water, soil, sediment and groundwater (Williams et al., 2016). Pharmaceuticals consist of a single or multiple active ingredients that have a specific biological activity. The therapeutic application and pharmacological mechanisms provide valuable information to evaluate the environmental hazard of these chemicals. Their physicochemical properties are of more relevance for the assessment of the environmental fate and exposure. The occurrence in the environment and the biological activity of this group of contaminants makes them relevant in environmental science.

References

Alder, A.C., Schaffner, C., Majewsky, M., Klasmeier, J., Fenner, K. (2010) Fate of [beta]-blocker human pharmaceuticals in surface water: Comparison of measured and simulated concentrations in the Glatt Valley Watershed, Switzerland. Water Research 44, 936-948

Aus der Beek, T., Weber, F., Bergmann, A., Hickmann, S., Ebert, I., Hein, A., Küster, A. (2016) Pharmaceuticals in the environment-Global occurrences and perspectives. Environ. Toxicol. Chem. 35, 823-835.

Guo, X.Y., Hao, L.J., Qiu, P.Z., Chen, R., Xu, J., Kong, X.J., Shan, Z.J., Wang, N. (2016) Pollution characteristics of 23 veterinary antibiotics in livestock manure and manure-amended soils in Jiangsu province, China. J. Environ. Sci. Health Part B Pestic. Food

Contamin. Agric. Wastes 51 (6), 383-392

Kaczala, F., Blum, S.E. (2016) The occurrence of veterinary pharmaceuticals in the environment: A review. Curr. Anal. Chem., 12 (3), 169-182

Kümmerer, K. (2009) The presence of pharmaceuticals in the environment due to human use - present knowledge and future challenges. J. Environ. Manage., 90 (8), 2354-2366

Monteiro, S.C., Boxall, A.B.A. (2010) Occurrence and Fate of Human Pharmaceuticals in the Environment. Rev. Environ. Contam. Toxicol., 202, 53-154.

Oosterhuis, M., Sacher, F., ter Laak, T.L. (2013) Prediction of concentration levels of metformin and other high consumption pharmaceuticals in wastewater and regional surface water based on sales data. Sci. Total Environ. 442, 380-388

Ter Laak, T.L., Van der Aa, M., Houtman, C.J., Stoks, P.G., Van Wezel, A.P. (2010) Relating environmental concentrations of pharmaceuticals to consumption: A mass balance approach for the river Rhine. Environ. Intern. 36, 403-409

Ter Laak, T.L., Kooij, P.J.F., Tolkamp, H., Hofman, J. (2014) Different compositions of pharmaceuticals in Dutch and Belgian rivers explained by consumption patterns and treatment efficiency. Environ. Sci. Pollut. Res. 21, 12843-12855

Van der Aa, N.G.F.M., Kommer, G.J., van Montfoort, J.E. and Versteegh, J.F.M. (2011) Demographic projections of future pharmaceutical consumption in the Netherlands. Water Science and Technology, 825-832.

Williams, M., Backhaus, T., Bowe, C., Choi, K., Connors, K., Hickmann, S., Hunter, W., Kookana, R., Marfil-Vega, R., Verslycke, T. (2016) Pharmaceuticals in the environment: An introduction to the ET&C special issue. Environ. Toxicol. Chem. 35, 763-766.

2.3.3. Question 1

Name three reasons why pharmaceuticals are relevant environmental contaminants?

2.3.3. Question 2

Why is it easier to model human pharmaceuticals in the environment than veterinary pharmaceuticals?

2.3.3. Question 3

Ibuprofen (anti-inflammatory) is discharged with the wastewater to the river. The following information is known about the use and fate of this drug.

- consumption per 1000 inhabitants of 1 g/d

- excretion by human 10 %

- removal by STP 68%

How much iboprufen is potentially discharged to the river in g/d ?

2.3.4. Drugs of abuse

Author: Pim de Voogt

Reviewer: John Parsons, Félix Hernández

Leaning objectives:

You should be able to:

- distinguish between licit and illicit drugs
- know what sources cause illicit drugs to show up in the environment
- understand what wastewater-based epidemiology is

Keywords: Cocaine, ecstasy, speed, cannabis, wastewater analysis

Introduction

Since about little more than a decade, drugs of abuse (DOA) and their degradation products have been recognized as emerging environmental contaminants. They are among the growing number of chemicals that can be observed in the aquatic environment.

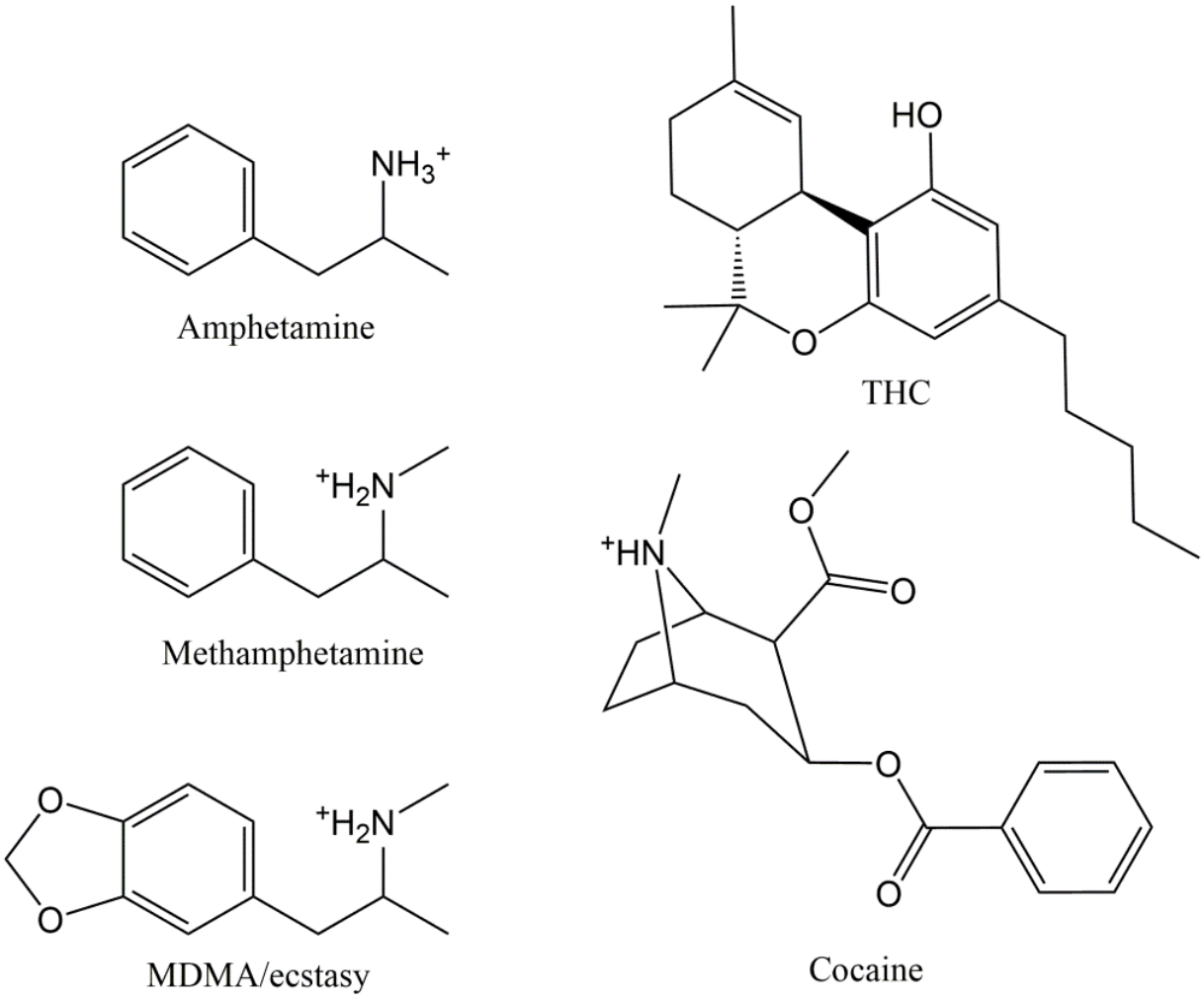


Figure 1.
Chemical structures of most popular drugs of abuse, in their predominant speciation under physiological conditions (pH7.4). (Source: Steven Droge)

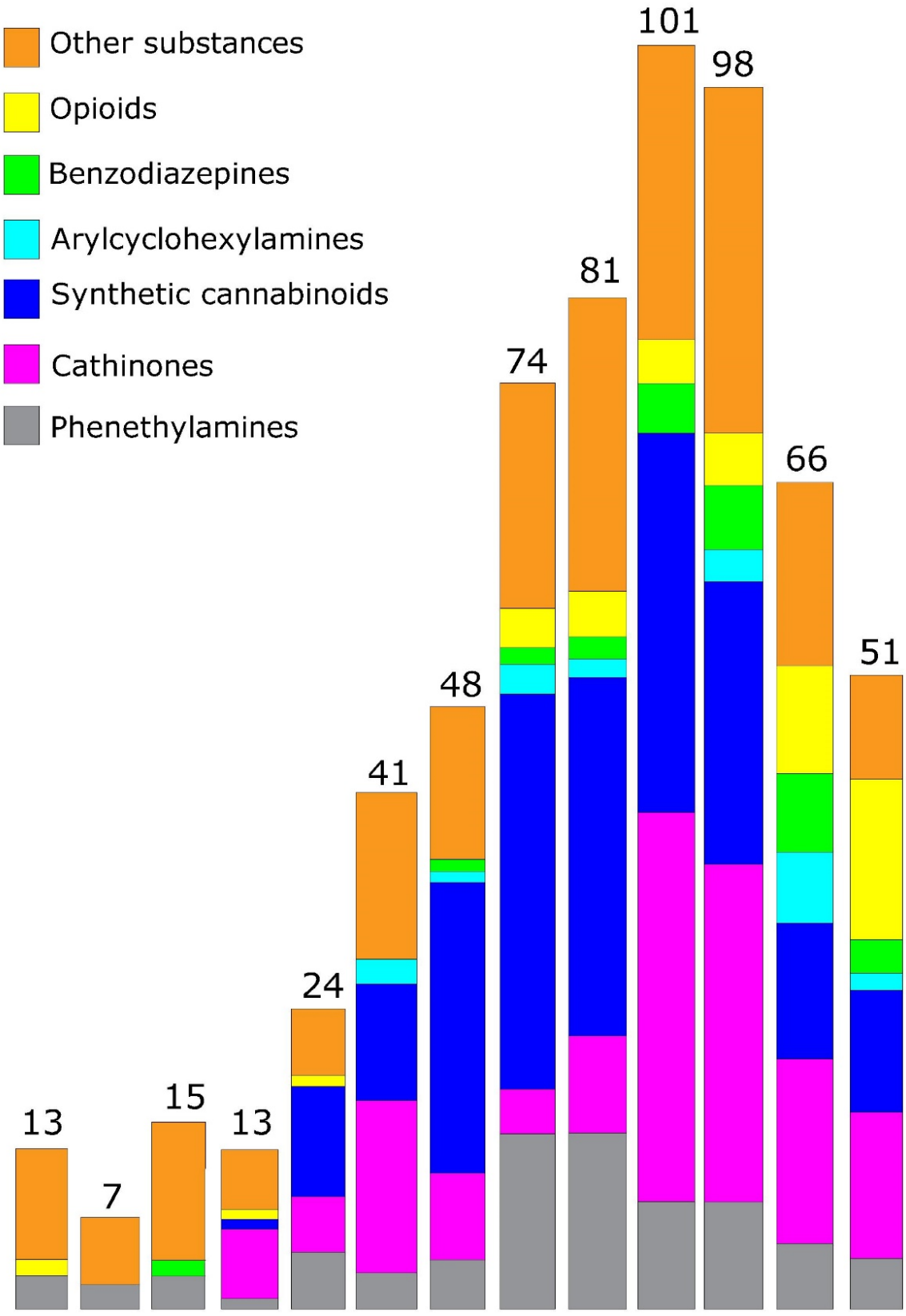
The residues of a major part of the chemicals used in households and daily life end up in our sewer systems. Among the many chemicals are cleaning agents and detergents, cosmetics, food additives and contaminants, pesticides, pharmaceuticals, and surely also illicit drugs. Once in the sewer, they are transported to wastewater treatment plants (WWTPs), where they may be removed by degradation or adsorption to sludge, or end up in the effluent of the plant when removal is incomplete.

The consumption of both pharmaceuticals and DOA has increased substantially over the last couple of decades as a result of several factors, including ageing of the population, medicalization of society and societal changes in life-style. As a result the loads in wastewater of drugs and their transformation products formed in the body after consumption have steadily increased. More recently, it has been observed that chemical waste from production sites of illicit drugs is being occasionally discharged into sewer systems, thereby dramatically increasing the loads of illicit drug synthesis chemicals and end products transported to WWTPs. As WWTPs are not designed to remove drugs, a substantial fraction of the loads may end up in receiving waters and thus pose a threat to both human and ecosystem health.

Drugs of Abuse (DOA)

Europe's most commonly used illicit drugs are THC (cannabis), cocaine, MDMA (ecstasy) and amphetamines. The structure of these drugs is given in Figure 1. Other important DOA include the opioids such as heroine and fentanyl, GHB, Khat and LSD.

Drugs of abuse are controlled by legislation, in The Netherlands by the Opium Act. The Opium Act encompasses two lists of substances. List one chemicals are called **hard drugs** while List II chemicals are known as **soft drugs**. Some narcotics are also being used for medicinal purposes, e.g., ketamine, diazepines, and one of the isomers of amphetamine. New psychoactive substances (NPS), also known as designer drugs or legal highs (because they are not yet controlled as they are not listed on the Opium Act lists), are synthesized every year and become available on the market in high numbers (see Figure 2).


Other substances
Opioids
Benzodiazepines
Arylcyclohexylamines
Synthetic cannabinoids
Cathinones
Phenethylamines
13
7
15
13
24
41
48
74
81
101
98
66
51

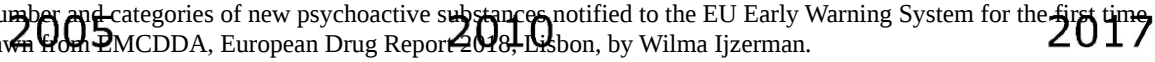


Figure 2. Number and categories of new psychoactive substances notified to the EU Early Warning System for the first time, 2005-2017. Redrawn from EMCDDA, European Drug Report 2018, Lisbon, by Wilma Ijzerman.

Wastewater-based epidemiology

Central sewage systems collect and pool wastewater from household cleaning and personal care activities as well as excretion products resulting from human consumption and thus contain chemical information on the type and amount of substances used by the population connected to the sewer. Drugs that are consumed are metabolized in the body and subsequently excreted. Excretion products can include the intact compounds as well as the transformation products, that can be used as biomarkers. An example of the latter is benzoylecgonine, which is the major transformation product of consumed cocaine. The collective wastewater from the sewer system carrying the load of chemicals is directed to the WWTP, and this wastewater influent can be sampled at the point where it enters the WWTP. By appropriate sampling of the influent during discrete time-intervals, e.g. 24 h, a so-called composite sample can be obtained and the concentrations of the chemicals can be determined. The volume of influent entering the WWTP is recorded continuously. Multiplying the observed 24 h average concentration of a compound with the total 24 h volume yields the daily load of the chemical entering the WWTP. This load can be normalised to the number of people living in the sewer catchment, resulting in a load per inhabitant. The loads of drugs in wastewater influents are usually expressed as $mg.day^{-1}.1000\ inh^{-1}$. Normalised drug load data allow comparison between sewer catchments, such as shown in Figure 3. Obtaining chemical information about the population through wastewater analysis is known as Wastewater-based epidemiology, WBE (*Watch the video*). While WBE was developed originally to obtain data on consumption of DOA, the methodology has been shown to have a much wider potential: in calculating the consumption of e.g., alcohol, nicotine, NPS, pharmaceuticals and doping, as well as for assessing community health indicators, such as incidence of diseases or stress biomarkers.

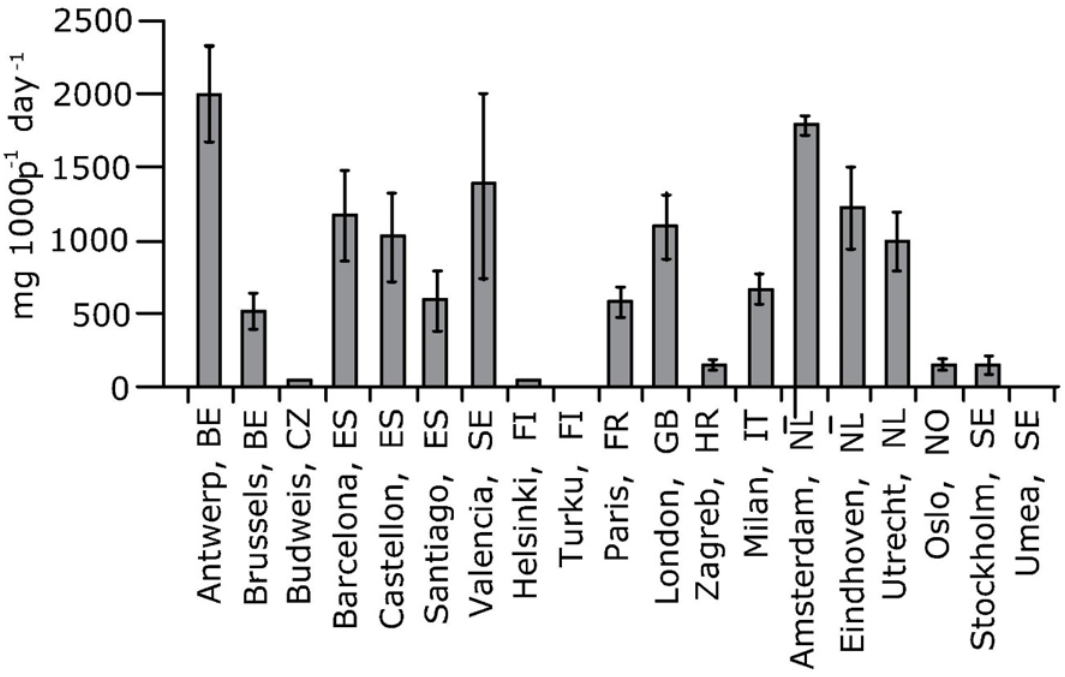


Figure 3.
Consumption of cocaine in 19 European cities in 2011 calculated from chemical analysis of influent loads of benzoylecgonine, a urinary biomarker of human cocaine consumption. Redrawn from Thomas et al. (2011) by Wilma Ijzerman.

DOA and the environment

Barring direct discharges into surface waters or terrestrial environments, the major sources of DOA to the environment are WWTP effluents. Conventional treatment in municipal WWTPs has not been specifically designed to remove pharmaceuticals or DOA. Removal rates of DOA vary widely and depend on compound properties such as persistence and polarity as well as WWTP operational conditions and process configurations. Some DOA cross WWTPs almost unhindered, thus ending op in the receiving waters. Examples of the latter are MDMA and some diazepines (see Figure 4). Despite that several studies report the presence of

DOA or their transformation products in surface waters, until now there is very little information about their aquatic ecotoxicity available in the scientific literature.

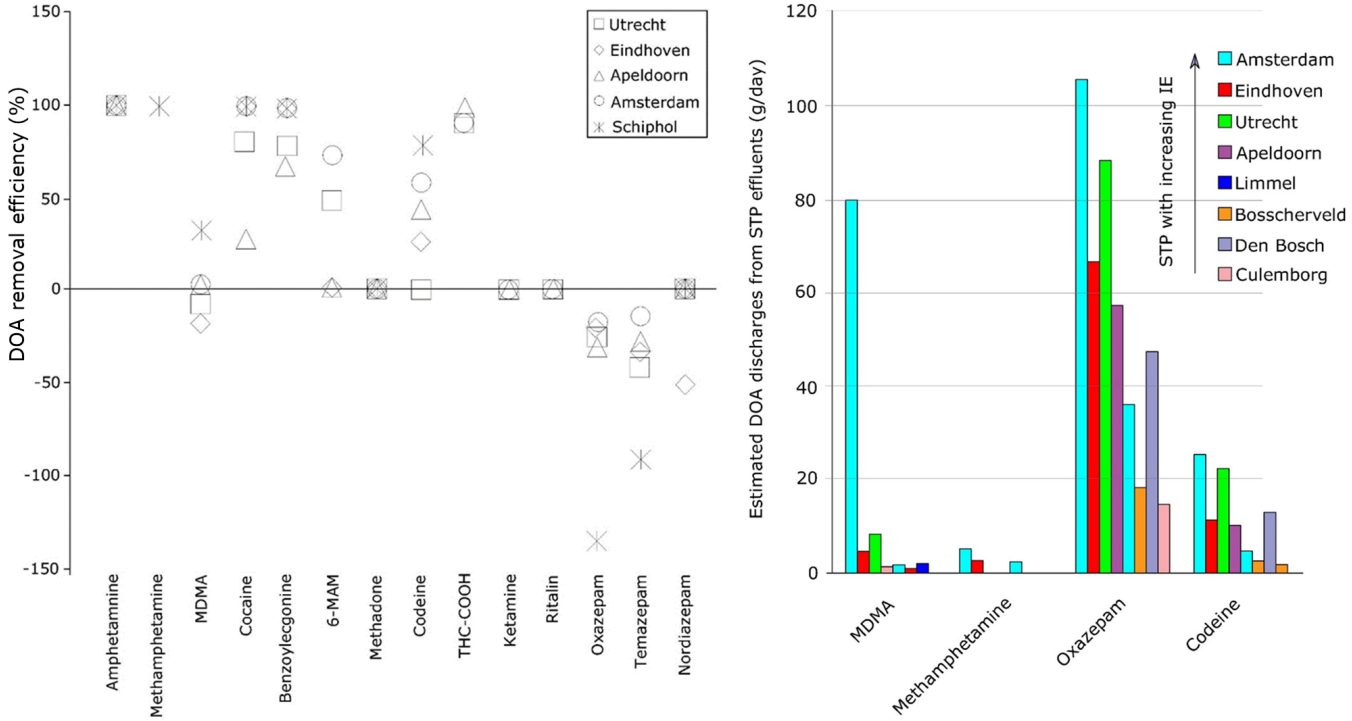


Figure 4. Data demonstrating that WWTP emit DOA to receiving waters. A) Removal efficiencies of DOA recorded in five Dutch WWTPs; B) Estimated discharges (g/day) of DOA from WWTPs based on monitoring data and WWTP effluent flow rates in 2009 (Sources: Bijlsma et al, 2012; Van der Aa et al, 2013). Drawn by Wilma IJzerman.

Recently, chemical waste from synthetic DOA manufacturing including their precursors and synthesis byproducts have been observed to be discharged directly into sewers. In addition, containers with chemical waste from DOA production sites have been dumped on soil or surface waters. Apart from solvents and acids or bases this waste often contains remainders of the synthesis products, which can then be dissipated in the aquatic environment or seep through the soil into groundwater.

Considering that DOA are highly active in the human body, it can be expected that some of them, in particular the more persistent ones, may exert some effects on aquatic biota when their levels increase in the aquatic environment.

References

Van der Aa, M., Bijlsma, L., Emke, E., et al. (2013). Risk assessment for drugs of abuse in the Dutch watercycle. Water Research 47(5), 1848-1857.

Bijlsma, L., Emke, E., Hernández, F., de Voogt, P. (2012). Investigation of drugs of abuse and relevant metabolites in Dutch sewage water by liquid chromatography coupled to high resolution mass spectrometry. Chemosphere 89(11), 1399-1406.

EMCDDA, European Drug Report 2018, Lisbon (www.emcdda.europa.eu/publications/edr/trends-developments/2018_en)

Thomas, K. V., Bijlsma, L., Castiglioni, S., et al. (2012). Comparing illicit drug use in 19 European cities through sewage analysis. Science of the Total Environment 432, 432-439.

2.3.4. Question 1

One way of mapping the consumption volumes of psychoactive substances in a city is by chemical analysis of drug residues and biomarkers in wastewater. To assess the consumption volume of cocaine by a population the concentrations in wastewater of the substance benzoylecgonine (BE) are determined. Why is BE used instead of cocaine itself?

2.3.4. Question 2

What possible sources can lead to the occurrence of an illicit drug in wastewater?

2.3.4. Question 3

Can illicit drugs in wastewater pose a threat to the environment and if so, why?

2.3.5. Hydrocarbons

Author: Pim N.H. Wassenaar

Reviewer: Emiel Rorije, Eric M.J. Verbruggen, Jonathan Martin

Learning objectives:
You should be able to

- explain the diversity/variation in hydrocarbon structures.
- explain the specific and non-specific toxicological effects of several hydrocarbons.

Keywords: Hydrocarbons, Paraffins, Naphthenics, Aromatics

Introduction:

Hydrocarbons are a class of chemicals that only consist of carbon and hydrogen atoms. But despite their simplicity in building blocks, this group of chemicals consists of a wide variety of structures, as there are differences in chain length, branching, bonding types and ring structures. The main sources of hydrocarbons are crude oil and coal, which are formed over millions of years by natural decomposition of the remains of plants, animals or wood, and are used to derive products we are using on a daily basis, including fuels and plastics Other natural sources include natural burning (forest fires) and volcanic sources

Hydrocarbon classification

The major classes of hydrocarbons are paraffins (i.e. alkanes), naphthenics (i.e. cycloalkanes) and aromatics (Figure 1), and within these classes, several subclasses can be identified. Paraffins are hydrocarbons that do not contain any ring structures. Paraffins can be subdivided in normal (n-) paraffins, which do not contain any branching (straight chain), and iso-paraffins (i-), which do contain a branched carbon-chain. When alkanes include at least one carbon-carbon double bond, they are considered olefins (or alkenes).

Naphthenic and aromatic hydrocarbons both contain ring-structures but differ in the presence of aromatic or non-aromatic rings. The naphthenics and aromatics can be further specified based on their ring count; often mono-, di- and poly-ring structures are distinguished from each other. Of all these classes, the polycyclic aromatic hydrocarbons (PAHs) are the best-studied category in terms of all kinds of environmental aspects.

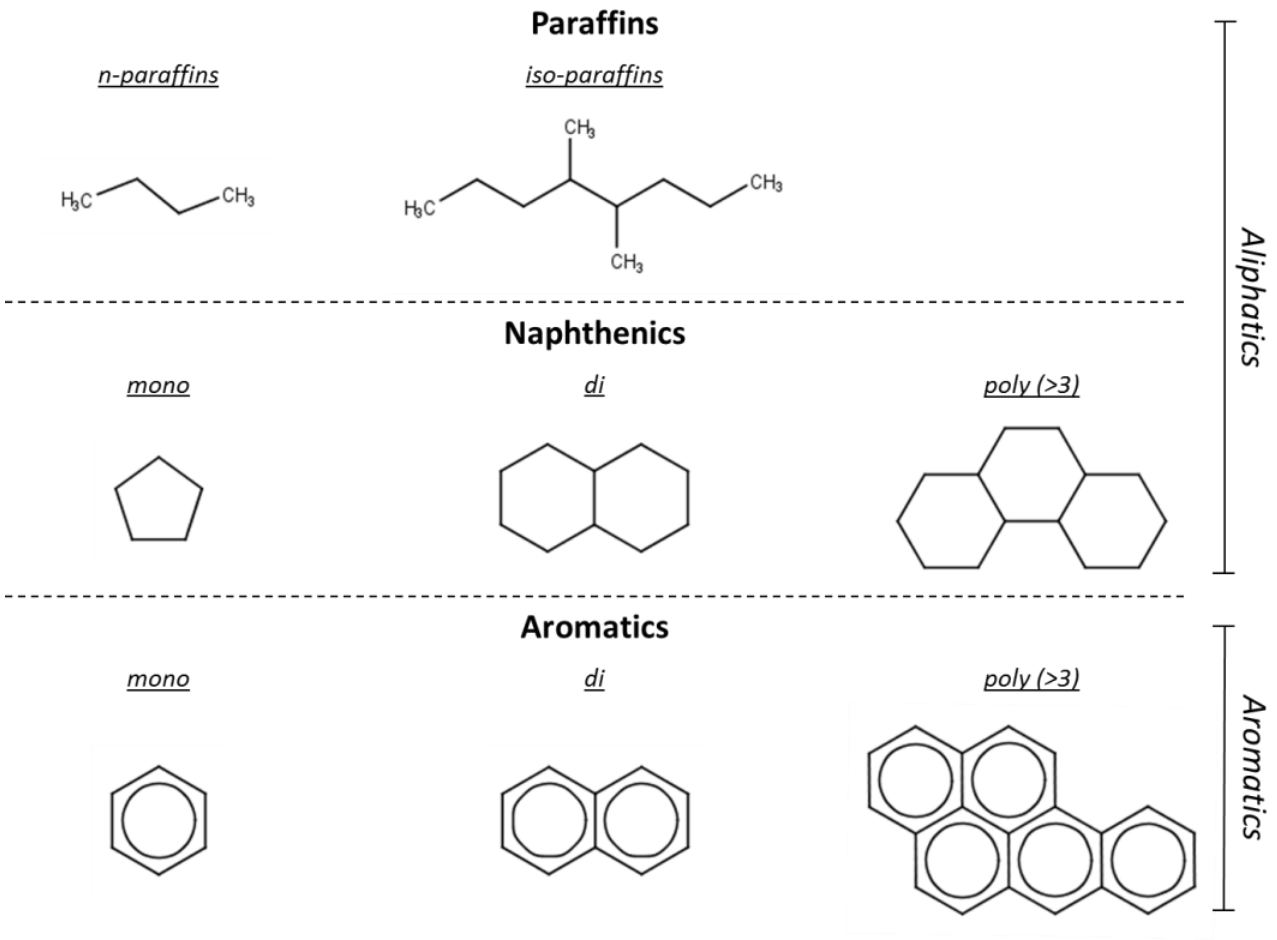


Figure 1. Chemical structures of common hydrocarbon classes. (by author)

Besides the classes considered in Figure 1., combinations of these classes also exist. Naphthenic or aromatic structures with an alkane side chain are mostly still considered as naphthenic or aromatic hydrocarbons, respectively. However, when a non-aromatic-ring is fused with an aromatic-ring, the hydrocarbon is classified as a naphthenic-aromatic structure. Depending on the ring-count several subclasses can be identified, including naphthenic-mono-aromatics and naphthenic-poly-aromatics.

Concerns for human health and the environment

Because of their lack of polar functional groups, hydrocarbons are generally hydrophobic and, as a consequence, many are able to cause acute toxic effects in aquatic animals by a non-specific mode of action known as narcosis (or baseline toxicity). Narcosis is a reversible state of inhibited activity of membrane structures within the cells of organisms. Narcosis type toxicity is considered the minimum toxicity that any substance will be able to have, just by reaching concentration levels in the phospholipid bilayer of the cell membranes that disturb membrane transportation process. Hence the name "baseline" or minimum toxicity. When these events take place above a certain threshold, systemic toxicity can be observed in the organism, such as lethality. This threshold concentration is also known as the critical body residue (CBR) (Bradbury et al., 1989; Parkerton et al., 2000; Veith & Broderius, 1990).

Nevertheless, hydrocarbons can also have a more specific mechanisms of action, resulting in greater toxicity than baseline toxicity. For example, the toxicity of several PAHs increases in combination with ultraviolet radiation due to photo-induced toxicity. Photo-induced toxicity may be caused by photoactivation, in which a PAH is degraded into an oxidized product with a higher toxicity, or rather by photosensitization, in which reactive oxygen species (ROS) are formed due to an excited state of the PAHs (Figure 2) (Roberts et al., 2017). PAHs are especially vulnerable to photodegradation as their absorption spectrum falls within the range of wavelengths reaching the earth's surface (> 290 nm), which is not the case for most monoaromatic and aliphatic hydrocarbons

(EMBSI, 2015). The photo-induced effects are of particular concern for aquatic species with transparent bodies, like zooplankton and early life stages, as more UV-light can penetrate into their organs and tissues (Roberts et al., 2017).

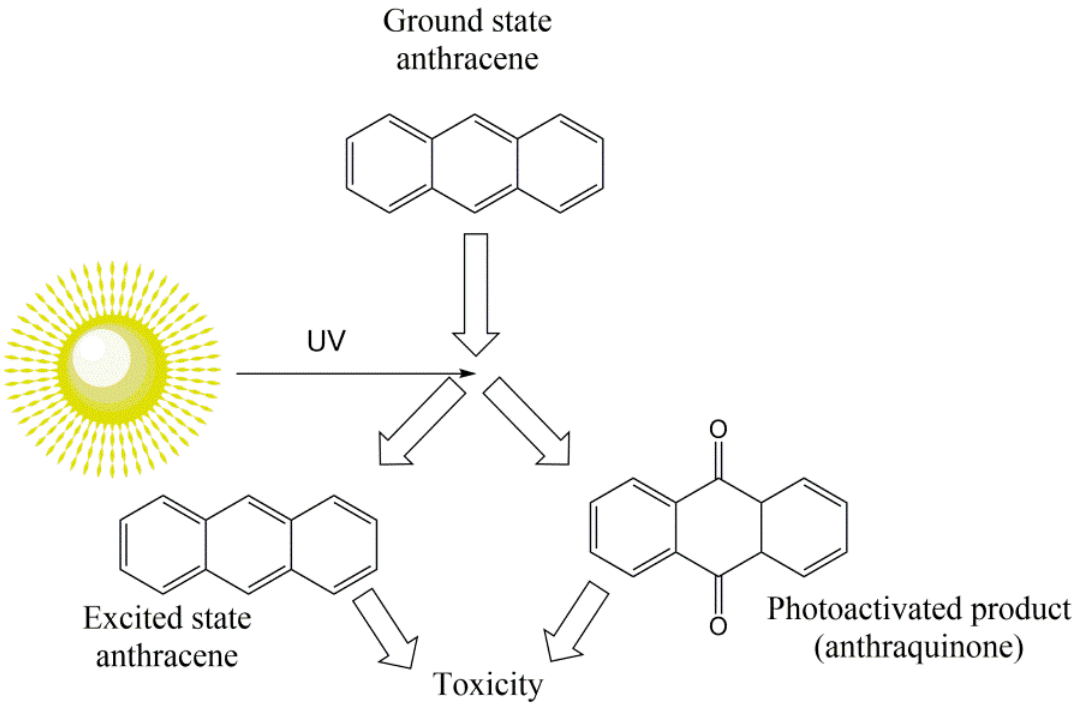


Figure 2. Mechanism of photo-induced toxicity of the polycyclic aromatic hydrocarbon anthracene via photosensitization or photomodification reactions, respectively. Adapted from Roberts et al. (2017) by Steven Droge.

Several hydrocarbons are also able to cause genotoxicity and cancer upon exposure, including benzene, 1,3-butadiene and some PAHs. The carcinogenicity of PAHs is caused by biotransformation into reactive metabolites, specifically into epoxides which are the first step in oxidation of aromatic ring structures into dihydrodiol ring systems (Figure 3). In general, the biotransformation step increases the water solubility of the hydrocarbons (Phase I metabolism) and promotes subsequent conjugation and excretion (Phase II metabolism). However, several epoxide metabolites - more specifically the most stable aromatic epoxides - can reach the cell nucleus and covalently react with DNA, forming DNA adducts, and induce mutations (Figure 3). Ultimately, if not repaired such mutations can accumulate and may result in the formation of tumors (Ewa & Danuta, 2016). Specifically, PAHs with a bay-like region are of concern as biotransformation results in relatively stable reactive epoxides that are not accessible to epoxide hydrolase enzymes (Figure 3) (Jerina et al. 1980). Similar to PAHs, 1,3-butadiene and benzene are also able to cause cancer via the effects of their respective reactive metabolites (Kirman et al., 2010; US-EPA 1998).

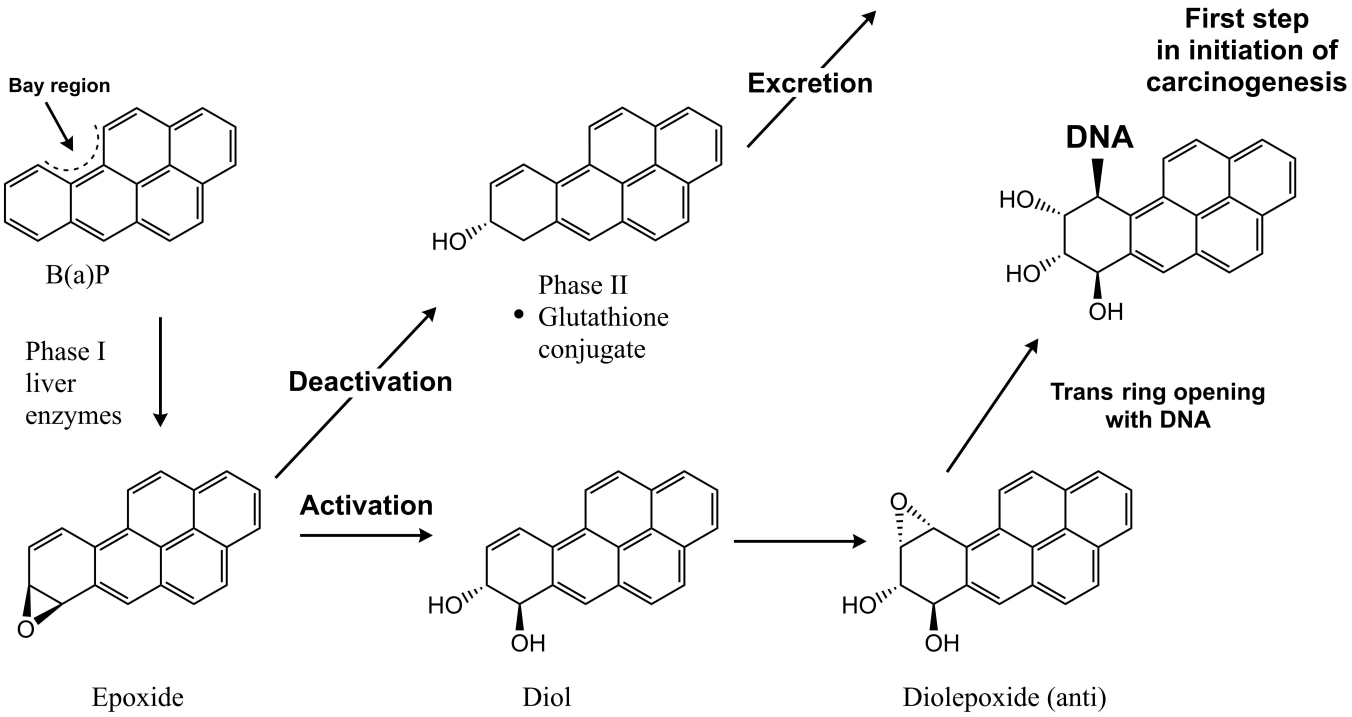


Figure 3. The biotransformation pathways of benzo(a)pyrene and binding to the DNA of reactive intermediates. Adapted from Homburger et al. (1983) by Steven Droge.

Besides their toxicity, some hydrocarbons such as the high molecular weight PAHs can be persistent in the environment and may accumulate in biota as a result of their hydrophobicity. It is therefore expected that internal concentrations are higher for such hydrocarbons and it is interesting that there is thus a relationship between narcosis and bioaccumulation potential. Consequently, these hydrocarbons might be of even greater concern.

Characterization of mixtures of hydrocarbons

As most research focused on specific hydrocarbons, including several PAHs, it is important to note that the biodegradation, bioaccumulation and toxicity potential of many hydrocarbons is still not fully known, such as for alkylated PAHs and naphthenics. As there is such a wide variety in hydrocarbon structures, it is impossible to assess the (potential) hazards of all hydrocarbons separately. Therefore, grouping approaches have been developed to speed up the risk assessment. Within a grouping approach, hydrocarbons can be clustered based on structural similarities. The underlying assumption is that all chemicals in a group are expected to have fairly similar physicochemical properties, and subsequently also fairly similar environmental fate and effect properties. As a result, such a group could potentially be assessed as if it is one single hydrocarbon.

The applicability of a hydrocarbon specific grouping approach, known as the Hydrocarbon Block Method (King et al., 1996), to assess the biodegradation and bioaccumulation potential of hydrocarbons is currently being investigated. Within this approach, all hydrocarbons are grouped based on their ***functional class*** (e.g. paraffin, naphthenic, aromatic) and the ***number of carbon atoms***. The number of carbon atoms is thought to highly correlate with the boiling point of the hydrocarbons. An example matrix of the Hydrocarbon Block Method is presented in Figure 4. The composition of an oil substance could be expressed in such a matrix following GC-GC/MS analysis. Subsequently, the PBT-properties of the individual blocks could potentially be assessed by analyzing and extrapolating the PBT-properties of representative hydrocarbons for varying hydrocarbon blocks (see Figure 4).

#C	Par	iP	MNap	DiNap	PolyNap	MAr	DiAr	PolyAr	NapMAr	NapDiAr	NapPolyAr	…	Total
…	…	…	…	…	…	…	…	…	…	…	…	…	…
11	3%	…	…	…	…	2%	…	…	…	…	…	…	5%
12	8%	4%	2%	1%	…	3%	1%	…	…	…	…	…	19%
13	15%	6%	4%	2%	…	4%	3%	1%	…	…	…	…	35%
14	9%	2%	2%	2%	…	4%	5%	3%	…	…	…	…	27%
15	1%	2%	1%	1%	…	3%	3%	2%	…	…	…	…	13%
16	…	…	1%	…	…	…	…	…	…	…	…	…	1%
17	…	…	…	…	…	…	…	…	…	…	…	…	…
18	…	…	…	…	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…	…	…	…	…
Total	36%	14%	10%	6%	…	16%	12%	6%	…	…	…	…	100%

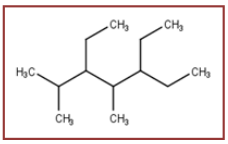
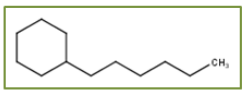
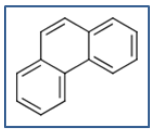




Figure 4. Theoretical example matrix of the hydrocarbon block method based on functional classes (columns) and carbon number (rows). Percentages represents the relative presence of specific hydrocarbon block within an oil substance. The PBT-properties of a block can potentially be assessed by analyzing and extrapolating PBT-properties of representative hydrocarbon structures.

References

Bradbury, S.P., Carlson, R.W., Henry, T R. (1989). Polar narcosis in aquatic organisms. In Aquatic Toxicology and Hazard Assessment: 12th Volume. ASTM International.

EMBSI (2015). Assessment of Photochemical Processes in Environmental Risk Assessment of PAHs

Ewa, B., Danuta, M.Š. (2017). Polycyclic aromatic hydrocarbons and PAH-related DNA adducts. Journal of applied genetics 58, 321-330.

Homburger, F., Hayes, J.A., Pelikanm E.W. (1983). A Guide to General Toxicology. Karger/Base, New York, NY.

Jerina, D.M., Sayer, J.M., Thakker, D.R., Yagi, H., Levin, W., Wood, A.W., Conney, A.H. (1980). Carcinogenicity of polycyclic aromatic hydrocarbons: the bay-region theory. In Carcinogenesis: Fundamental Mechanisms and Environmental Effects (pp. 1-12). Springer, Dordrecht.

King, D.J., Lyne, R.L., Girling, A., Peterson, D.R., Stephenson, R., Short, D. (1996). Environmental risk assessment of petroleum substances: the hydrocarbon block method. CONCAWE report no. 96/52.

Kirman, C.R., Albertini, R.A., & Gargas, M.L. (2010). 1, 3-Butadiene: III. Assessing carcinogenic modes of action. Critical reviews in toxicology 40(sup1), 74-92.

Parkerton, T.F., Stone, M.A., Letinski, D. J. (2000). Assessing the aquatic toxicity of complex hydrocarbon mixtures using solid phase microextraction. Toxicology letters 112, 273-282.

Roberts, A.P., Alloy, M.M., Oris, J.T. (2017). Review of the photo-induced toxicity of environmental contaminants. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 191, 160-167.

US-EPA (1998). Carcinogenic Effects of Benzene: An Update. EPA/600/P-97/001F.

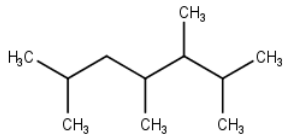
Veith, G.D., Broderius, S.J. (1990). Rules for distinguishing toxicants that cause type I and type II narcosis syndromes. Environmental Health Perspectives 87, 207.

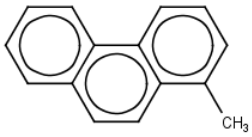
2.3.5. Question 1

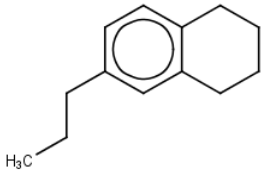
Mention at least three aspects that influence the variation in hydrocarbon structures?

2.3.5. Question 2

To which hydrocarbon class (or Hydrocarbon Block) do the following structures belong?

1. 

2. 

3. 

2.3.5. Question 3

What kind of toxic effects can be observed for polycyclic aromatic hydrocarbons?

2.3.6. CFCs

(draft)

Authors: Steven Droge

Reviewer: John Parsons

Leaning objectives:

You should be able to:

- realize what the ozone layer depletion was all about
- understand why certain replacement chemicals are still problematic

Keywords: Ozone layer, refrigerator, volatile chemicals, spray cans, radicals

Introduction

CFCs (chlorofluorocarbons) were very common air pollutants in the 20th century because they were the basic components of refrigerants and air conditioning, propellants (in spray can applications), and solvents, since the 1930s. They are still very common air pollutants, because they are very persistent chemicals, and emissions do still continue. In the first years as refrigerants, they replaced the much more toxic components ammonia (NH_3), chloromethane (CH_3Cl), and sulfur dioxide (SO_2). Particularly the CFCs leaking from old refrigerating systems in landfills and waste disposal sites caused high emissions into the environment. Typically, these volatile CFC chemicals are based on the smallest carbon molecules methane (CH_4), ethane (C_2H_6), or propane (C_3H_8). All hydrogen atoms in these CFC molecules are replaced by a mixture of chlorine and fluorine atoms.

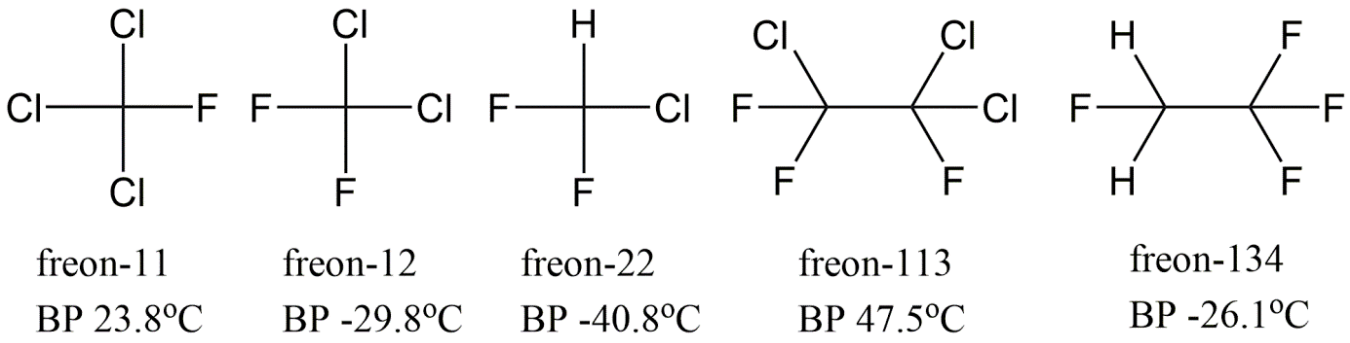


Figure 1. Different common refrigerants and their boiling points. Freon 134 is chlorine free. . (Source: Steven Droge)

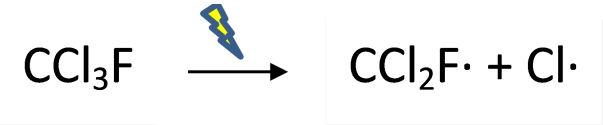
CFCs are less volatile than their hydrocarbon analogue, because the halogen atoms polarize the molecules, which causes stronger intermolecular attractions. Depending on the substitution with Cl or F, the boiling point can be tuned to the desired point for refrigerating cooling processes. The CFCs are also much less flammable than hydrocarbon analogues, making them much safer in all kinds of applications.

Naming of CFCs

CFCs were often known by the popular brand name Freon. Freon-12 (or R-12) for example stands for dichlorodifluoromethane (CCl_2F_2, boiling point -29.8 °C, while methane has -161 °C), as shown in Figure 1. The naming reflects the amount of fluor atoms as the most right number. The next value to the left is the number of hydrogen atoms plus 1, and the next value to the left is the number of carbon atoms less one (zeroes are not stated), and the remaining atoms are chlorine. Accordingly, Freon-113 could apply to 1,1,2-trichloro-1,2,2-trifluoroethane ($C_2Cl_3F_3$, boiling point 47.7 °C, while ethane has -161 °C). The structure of any Freon-X number can also be derived from adding +90 to the value of X, so Freon-113 would give a value of 203. The first numerical is the number of C (2), the second numerical H (0), the third numerical F (3), and the remaining substitutions are by chlorine (C_2X_6 gives 3 chlorines).

The reason CFC depletes the ozone layer

The key issue with CFC emissions is the reaction under influence of light ("photodegradation") that ultimately reduces ozone concentrations ("ozone depletion") in the upper atmosphere ("stratosphere"). Ozone absorbs the high energy radiation of the solar UV-B spectrum (280-315nm), and the ozone layer therefore prevents this to reach the Earth's surface. The even more energetic solar UV-C spectrum (100-280nm) is actually causing the formation of ozone (O_3) when reacting with oxygen (O_2), as shown in Figure 2. Under the influence of intense light-energy in the upper atmosphere, CFC molecules can disintegrate into two highly reactive radicals (molecules with a free electron ·), for Freon-11:



It is the radical Cl· that catalyzes the conversion of ozone back into O_2. The environmentally relevant role of the fluorine atoms in CFCs is that they make these chemicals very persistent after emission, because the C-F bond is one of the strongest covalent bonds known. With half-lives up to >100 years, high CFC levels can reach the upper atmosphere. James Lovelock was the first to detect the widespread presence of CFCs in air in the 1960s, while the damage caused by CFCs was discovered only in 1974. Another undesirable effect of CFC in the stratosphere is that they are a much more potent greenhouse gases than CO_2.

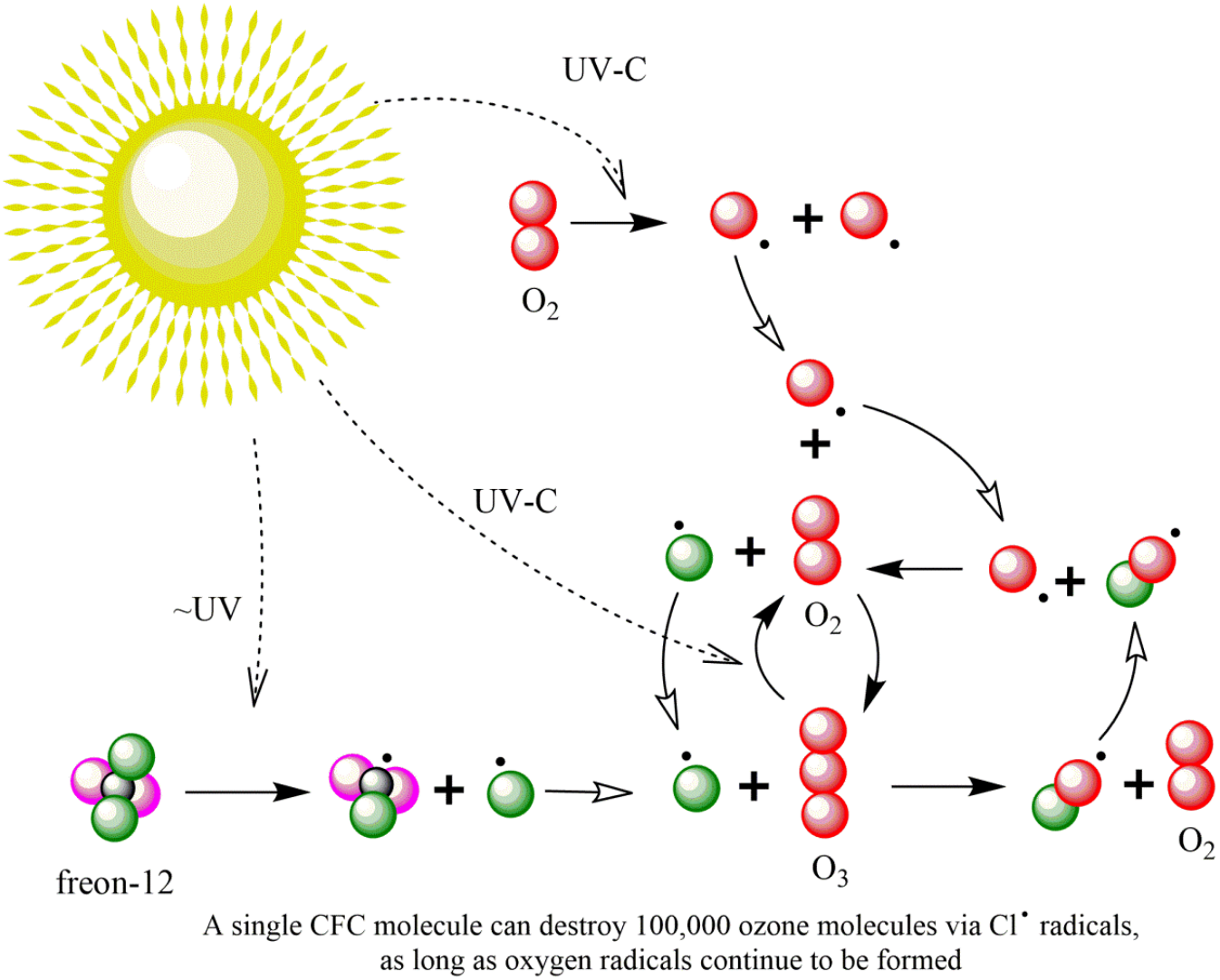


Figure 2. The influence of UV on formation of chlorine radicals from CFCs, oxygen radicals from O_2 and oxygen radicals from disintegration of O_3. Ozone is not formed in the absence of UV (night time), but can still be reacting away by chlorine radicals. (Source: Steven Droge)

CFC replacements.

In 1978 the United States banned the use of CFCs such as Freon in aerosol cans. After several years of observations of the ozone layer depletions globally (Figure 3), particularly above Antarctica, the Montreal Protocol was signed in 1987 to drastically reduce CFC emissions worldwide. CFCs were banned by the late 1990s in most EU countries, and e.g. in South Korea by 2010. Due to the persistency of CFCs it may take until 2050-2070 before the ozone layer will return to 1980 levels (which were bad already).

The key damaging feature of CFCs in terms of ozone depletion is their persistency, so that emissions reach and build up in the stratosphere (starting from 20km above the equator, but only at 7km above the poles). CFC replacement molecules were initially found simply by adding more hydrogens in the CFC structures and somewhat less Cl (HCFCs), but fractions still contributed to Cl· radicals. Later alternatives lack the chlorine atoms and have even shorter lifetimes in the lower atmosphere, and simply cannot form the Cl radicals. These "hydrofluorocarbons" (HFCs) are currently common in automobile air conditioners, such as Freon-134 (do the math to see that there is no Cl, boiling point -26.1 °C).

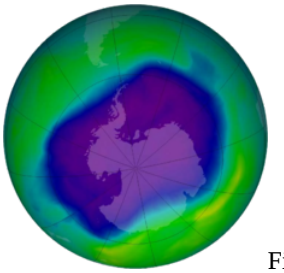


Figure 3. The 2006 record size hole in the ozone layer above Antarctica (Source https://en.Wikipedia.org/wiki/Ozone_depletion)

Still, HCFC as well as HFCs are still very potent greenhouse gasses, so the worldwide use of such chemicals remains problematic and gives rise to new legislations, regulations, and searches for alternatives. R-410A (which contains only fluorine) is becoming more widely used but is 1700 times more potent than CO_2 as greenhouse gas, equal to Freon-22. Simple hydrocarbon mixtures such as propane/isobutane are already used extensively in mobile air conditioning systems, as they have the right thermodynamic properties for some uses and are relatively safe. Unfortunately, we did not have the technological skills, nor the awareness to apply this back in the 1930s.

2.3.7. Cosmetics/personal care products

(draft)

Author: Mélanie Douziech

Reviewers: John Parsons

Learning objectives:

You should be able to:

- Define what personal care products and cosmetics are
- Explain how chemicals from personal care products end up in the environment
- Cite and describe some of the most common chemicals found in personal care products

Keywords: wastewater, chemical function, surfactants, microbeads

Introduction

Personal care products (PCPs) cover a large range of products fulfilling hygiene, health, or beauty purposes (e.g. shampoo, toothpaste, nail polish). They are categorized into oral care, skin care, sun care, hair care, decorative cosmetics, body care and perfumes. Overall, most PCPs are classified as cosmetics and regulated accordingly. In the European Union (EU) the Cosmetic Regulation governs the production, safety of ingredients and the labelling and marketing of cosmetic products. The United States of America (USA), on the other hand, have a narrower definition of cosmetics so that products not fulfilling the definition are regulated as pharmaceuticals (e.g. sunscreen) (Food and Drug Administration, 2016).

PCPs come in a range of formats (e.g. liquids, bars, aerosols, powders) and typically contain a wide range of chemicals, each fulfilling a specific function within the product. For example, a shampoo can include cleansing agents (surfactants), chemicals to ensure product stability (e.g. preservatives, pH adjusters, viscosity controlling agents), diluent (e.g. water), perfuming chemicals (fragrances), and chemicals to influence the product's appearance (e.g. colourants, pearlescers, opacifiers). The chemicals present in PCPs ultimately enter the environment either through air during direct use, such as the propellants in aerosols, or through wastewater via down the drain disposal following product use (e.g. shower products, toothpaste). The release of PCP chemicals into the environment needs to be monitored and the safety of these chemicals understood in order to avoid potential problems. In developed countries, the use of wastewater treatment plants (WWTPs) is key to effectively removing the PCP chemicals and other pollutants from wastewater prior to their release to rivers and other watercourses. The removal mechanisms occurring in WWTPs include biodegradation, sorption onto solids, and volatilization to the air. The extent of removal is influenced by the physicochemical properties of the chemicals and the operational conditions of the WWTPs. In regions where wastewater treatment is lacking, the chemicals in PCPs enter the environment directly.

The wide scale daily use of PCPs and the associated large volumes of chemicals released explain why they are scrutinized by environmental protection agencies and regulatory bodies. The following sections will briefly review some of the classes of chemicals used in PCPs by describing their behavior in the environment and their potential effect on ecosystems.

Cleansing agents - surfactants

Surfactants are an important and widely used class of chemicals. They are the key components of many household cleaning agents as well as PCPs, such as shampoos, soaps, bodywash and toothpaste, because of their ability to remove dirt. These dirt-removing properties also make surfactants inherently toxic to aquatic organisms. The biodegradability of surfactants is a key legal requirement for their use in PCPs to minimize the likelihood of unsafe levels in the environment. Different types of surfactants exist and are often classified based on their surface charge. Anionic surfactants, which carry a negative surface charge, interact and help remove positively charged dirt particles from surfaces such as hair and skin. Sodium lauryl sulfate is a typical example of an anionic surfactant used in PCPs. Cationic surfactants, such as cetrimonium chloride, are positively charged and may be used as hair conditioning agents to make hair shinier or more manageable. Non-ionic surfactants (uncharged), such as cetyl alcohol, help formulate products or increase foaming. Amphoteric surfactants, such as sodium lauriaminodipropionate, carry both positive and negative charges and are commonly used to counterbalance the potentially irritating properties of anionic surfactants.

Fragrances

Fragrances are mixtures of often more than 20 perfumery chemicals used to provide the smell of PCPs. Typically, fragrances are present at very low levels in most PCPs (below 0.01%) so that their exact compositions are not disclosed. Disclosed, however, are any allergens present in the fragrance to help dermatologists and consumers avoid certain fragrance chemicals. Despite the wish to protect trade secrets, a recent trend increasingly sees companies disclose the full fragrance compositions of their products on their websites (e.g. L'Oréal, Unilever). Well-known examples of fragrances include hexyl cinnamal, linalool, and limonene. Potential concerns about the ecotoxicological impact of fragrances have arisen on the one hand because of a lack of disclosure of fragrance formulations and on the other hand because of the detection of certain persistent fragrances in the environment (e.g. nitromusks).

Preservatives

Preservatives are usually added to PCPs containing water for their ability to protect the product from contamination by bacteria, yeasts, and molds during storage or repeated use. Given their targeted action against living organisms, the use of preservative in chemical products including PCPs is under constant scrutiny. For example, in 2016 and 2017, the European Commission tightened the regulation around the use of methylisothiazolinone in cosmetics products due to human safety concerns. Other preservatives that have been restricted in use, because of both human safety and environmental safety concerns (e.g. endocrine disruption effects), include certain types of parabens and triclosan.

UV filters

UV filters are used in sunscreen products as well as in other PCPs such as foundation, lipstick, or moisturizing cream to protect users from UV radiation. UV filters can be organic or inorganic. Inorganic UV filters, like titanium oxide and zinc oxide, form a physical boundary protecting the skin from UV radiation. Organic UV filters, on the other hand, protect the skin by undergoing a chemical reaction with the incoming UV radiation. Organic UV filters commonly found in PCPs include butyl methoxydibenzoylmethane, ethylhexyl methoxycinnamate, and octocrylene. Organic UV filters are poorly biodegradable and have the potential to accumulate in organisms. Further, a number of organic UV filters have been shown to be toxic to coral organisms in laboratory tests. They are suspected to cause coral bleaching by, for example, promoting viral infections but research is still on-going to understand their potential ecotoxicological effects at realistic environmental concentrations.

Volatile chemicals

Certain chemicals used in PCPs are highly volatile and may end up in the air following product use. Examples include propellants, such as propane butane mixes or compressed air/nitrogen, used in aerosols to apply ingredients in hairsprays or deodorants and antiperspirants. Fragrances also volatilize when the product is applied to skin or hair to provide smell. Volatile silicones, chemicals used to assist the deposition of ingredients in liquids and creams, are another example of chemicals emitted to air upon PCP use.

The special case of plastic microbeads

Plastic microbeads, with a diameter smaller than 5mm, have been used in PCPs such as face scrubs or shower gels for their scrubbing and cleansing properties. The growing concern about plastic pollution in water has drawn attention to the use of microbeads in PCPs. As a result, a number of initiatives were launched both to highlight the use of plastic microbeads and to encourage replacement with natural alternatives. An example thereof is the "Beat the microbead" coalition (https://www.beatthemicrobead.org/) sponsored by the United Nations Environment Program, launched to help consumers identify and avoid PCPs containing microbeads. Such initiatives together with voluntary commitments by industry have led to a large decrease in the use of microbeads in wash-off cosmetic products: In the EU, for example, the use of microbeads in wash-off products was reduced by 97% from 2012 to 2017. Legislation to restrict the use of microbeads has also recently been put in place. In the USA microbeads in PCPs were banned in July 2017 and a number of EU countries (e.g. United Kingdom, Italy) have also banned their use in wash-off products.

Further reading

For more information on PCP chemicals and their function in products, please see (European commission 2009; Grocery Manufacturers Association 2017).

For more information on the different types of surfactants, please see Tolls et al. (2009) and Section 2.3.8.

Manova et al. (2013) list the different types of UV filters.

The report of Scudo et al. (2017) gives more information on the use of microplastics in Europe.

References

European Commission (2019). Cosing. 2009 03.2019]; Available from: http://ec.europa.eu/growth/tools-databases/cosing/.

Food and Drug Administration (2016). Are All "Personal Care Products" Regulated as Cosmetics? [cited 2019 03]; Available from: https://www.fda.gov/forindustry/fdabasicsforindustry/ucm238796.htm.

Grocery Manufacturers Association (2017). Smartlabel. [cited 2017 11]; Available from: http://www.smartlabel.org/.

Manova, E., von Goetz, N., Hauri, U., Bogdal, C., Hungerbuhler, K. (2013). Organic UV filters in Personal Care Products in Switzerland: A Survey of Occurrence and Concentrations. International Journal of Hygiene and Environmental Health 216, 508-514.

Scudo, A., Liebmann, B., Corden, C., Tyrer, D., Kreissig, J., Warwick, O. (2017). Intentionally Added Microplastics in Products. in: Limited A.F.W.E.a.I.U., ed. United Kingdom

Tolls, J., Berger, H., Klenk, A., Meyberg, M., Müller, R., Rettinger, K., Steber, J. (2009). Environmental safety aspects of Personal Care Products - a European perspective. Environmental Toxicology and Chemistry 28, 2485-2489.

2.3.7. Question 1

How do chemicals found in cosmetics end in the environment?

2.3.7. Question 2

Why are surfactants a concern for environmental toxicity?

2.3.7. Question 3

How did personal care products contribute to the microplastic pollution in the water?

2.3.8. Detergents and surfactants

Author: Steven Droge

Reviewer: Thomas P. Knepper

Leaning objectives:

You should be able to:

- explain why surfactants remove dirt
- discuss historical progress on surfactant biodegradability
- describe the different types of common surfactants.
- describe examples of how surfactants enter the environment.

Keywords: amphiphilic chemicals, micelle formation, biodegradability

Introduction

Surface active agents ("surf-act-ants") are a wide variety of chemicals produced in bulk volumes (>10.000 tonnes annually) as a key ingredient in cleaning products: detergents. Typical for surfactants is that they have a hydrophobic tail and a hydrophilic head group (Figure 1).

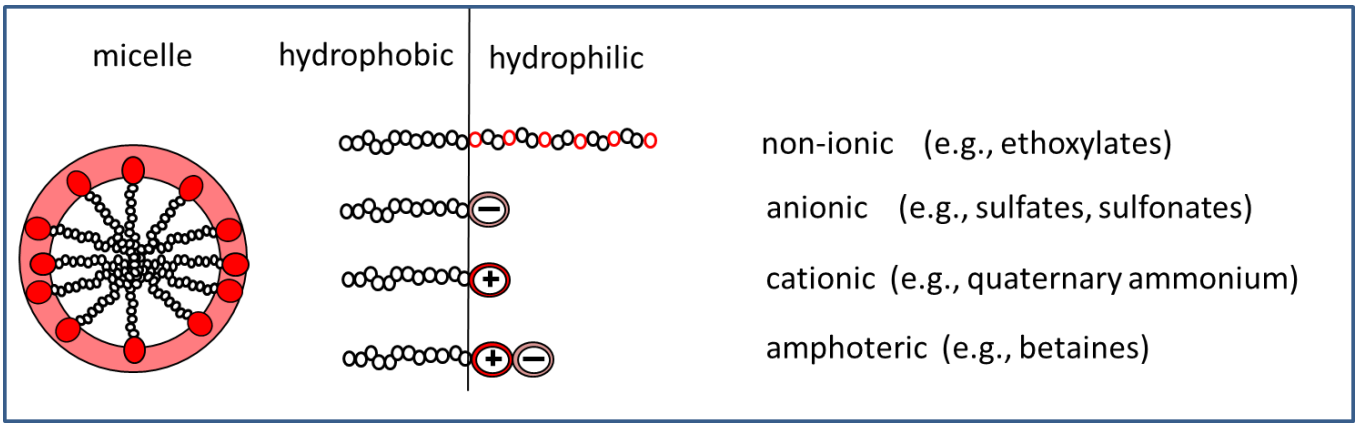


Figure 1. Different forms (micelle and surfactant monomer) and types of surfactants. (Source: Steven Droge)

At relatively high concentrations in water (typically >10-100 mg/L), surfactants spontaneously form aggregated structures called micelles (Figure 1), often in spheres with the hydrophobic tails inward and the hydrophilic head groups towards the surrounding water molecules. These micelle super-structures allow surfactants to dissolve grease and dirt from e.g. textile or dishes into water, which can then be flushed away. Besides this common use of surfactants, their amphiphilic (i.e., both hydrophilic and lipophilic) properties allow for a versatile use in our modern world:

• During the large 2010 oil spill in the Mexican Gulf, enormous volumes (>6700 tonnes) of several types of surfactant formulations (e.g. "Corexit") were used to disperse the constant stream of oil leaking from the damaged deep water well into small dissolved droplets, in order to facilitate microbial degradation and prevent the formation of floating oil slabs that could ruin coastal habitats.

• The ability of a layer of surfactants to maintain hydrophobic particles in solution is a key process in many products, such as paints and lacquers.

• The ability to emulsify dirt particles is a key feature in process fluids during deep drilling in soil or sediment.

• Fabric softners, and hair conditioners, have cationic surfactants as key ingredients that stick with the positively charged head groups onto the negatively charged fibers of your towel or hair. After the final flushing, these cationic surfactants still stick on the fibers and because of the hydrophobic head groups sticking out make these materials feel soft and smooth. Often only during the next washing event (with anionic or nonionic surfactants) the cationic surfactants are flushed off the fibers.

• Many cationic surfactants have biocidal properties at relatively low concentrations and are therefore used in a few percent in many cosmetic products as preservatives, e.g. in cosmetics, or used to kill microbes in food processing, antibacterial hand wipes or during swimming pool cleaning. Examples are chloride salts of benzalkonium, benzethonium, cetylpyridinium.

• Surfactants lower the surface tension of water, and therefore are used (as "adjuvants") in pesticide products to facilitate the droplet formation during spraying and to improve contact of the droplets with the target leaves in case of herbicides. Examples are fluorinated surfactants, silicone based surfactants (Czajka et al. 2015), and polyethoxylated tallow amine (POEA) used for example in the glyphosate formulation Roundup.

The hydrophobic tail of surfactants is mostly composed of a chain of carbon atoms, although also fluorinated carbon (-CF_2-) chains or siloxane ($Si(CH_3)_3$-O-[..]-$Si(CH_3)_3$) chains are also possible.

The first bulk volume produced surfactants for washing machines were **branched** anionic alkylbenzenesulfonates (ABS) and alkylphenolethoxylates (APEO), with the hydrocarbon source obtained from petroleum. Because of the variable petroleum source, these chemicals are often complex mixtures. However, hydrophobic branched alkylchains are poorly biodegraded, and the constant disposal of these surfactants into the waste water caused very high environmental concentrations, often leading to foaming rivers (Figure 2).

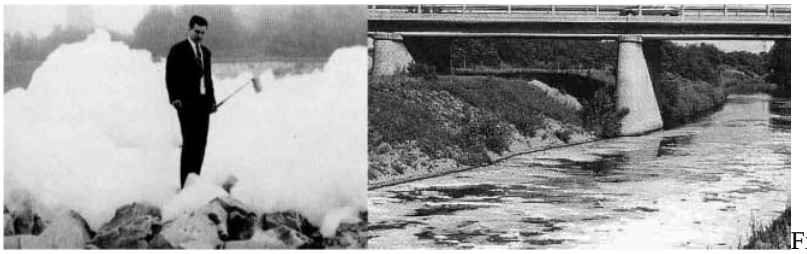


Figure 2. Foaming in sewage treatment and on rivers in the 1950s caused by non-biodegradable tetrapropylene sulfonate (TPS) (Source: Kümmerer 2007, who obtained permission from the Archives of the Henkel Company, Düsseldorf).

Surfactant producers 'voluntarily' switched to carbon sources such as palm oil or controlled polymerization of petrol-based ethylene, that could be used to generate surfactants with **linear alkyl chains**: linear alkylbenzenesulfonate (LAS) and alcohol ethoxylates (AEO). Some surfactants have the hydrophilic headgroup attached to two carbon chains, such as the anionic docusate (heavily used in the BP oil spill) and the cationic dialkyldimethylammonium chemicals. Common detergent surfactants are nowadays designed to pass ready biodegradability tests (>60% mineralisation to CO_2 within a 10 d window following a lag phase, in a 28 d test). Early examples of fabric softners are double chain (dialkyl)dimethylammonium surfactants, but the environmental persistency of these compounds (DODMAC and DHTDMAC, see e.g. EU and RIVM-reports) has led to a large replacement by diesterquats (DEEDMAC), which degrade more rapidly through the weak ester linkages of the fatty acid chains (Giolando et al. 1995). A switch to sustainable production of the carbon sources is ongoing. Whereas petroleum based ethylene oil was mostly used, it is being replaced increasingly by the linear fatty acid carbon chains from either palm-oil (mostly C_{16}/C_{18}), coconu oil (mostly C_{12}/C_{14}), but also such raw materials needs to be as sustainably derived as possible.

The hydrophilic headgroups can vary extensively. Nonionic surfactants can have a simple polar functional group (amide), glucose based (polyglycoside), or contain a variable lengths of repetitive ethoxlyate and/or propoxylate units. Because the ethoxylation process is difficult to control, such surfactants are often complex mixtures. Anionic surfactants are often based on sulfate (SO_4-) or sulfonate (SO_4-), but also phosphonate and carboxylates are common. A key difference between anionic surfactants is that sulfate and sulfonates are fully anionic (pKa ~<0) over the entire environmental pH range (pH4-9), while carboxylates are weaker acids that are still partially neutral species (pKa ~5). Most cationic surfactants are based on permanently charged quaternary ammonium headgroups (R-(N+)$(CH_3)_3$), although several ionizable amine groups are applied in cationic surfactants too (e.g., diethanolamines).

The key ingredient property of most surfactants is the critical micelle concentration (CMC), which defines the dissolved concentration above which micellar aggregates start to form that can remove grease or fully emulsify particles. The CMC decreases proportionally with the hydrophobic tail length, and this means that with longer tails, you need less surfactant to start to form micelles. However, with increasing hydrophobic tails anionic surfactants more readily precipitate with dissolved inorganic cations such as calcium. Also, surfactant toxicity increases proportionally with hydrophobic tail lengths. If the alkyl chain is too long, the surfactant may bind strongly to all kinds of surfaces and not be available for micelle formation. The optimum hydrophobic chain length is thus often a balance between the desired properties of the surfactant and several critical processes that influence the efficiency and risk of surfactants.

References

Kümmerer, K. (2007). Sustainable from the very beginning: rational design of molecules by life cycle engineering as an important approach for green pharmacy and green chemistry. Green Chemistry 9, 899-907 . DOI: 10.1039/b618298b

Giolando, S.T., Rapaport, R.A. , Larson, R.J., Federle, T.W., Stalmans, M., Masscheleyn P. (1995). Environmental fate and effects of DEEDMAC: A new rapidly biodegradable cationic surfactant for use in fabric softeners. Chemosphere 30, 1067-1083. DOI: 10.1016/0045-6535(95)00005-S

Czajka, A., Hazell, G., Eastoe, J. (2015). Surfactants at the Design Limit. *Langmuir* 31, 8205−8217. DOI: 10.1021/acs.langmuir.5b00336

Scientific Committee on Toxicity, Toxicity and the Environment (CSTEE) (2001). Opinion on the results of the Risk Assessment of: Dimethyldioctadecylammonium chloride (DODMAC). EU-report C2/JCD/csteeop/DodmacHH22022002/D(02) https://ec.europa.eu/health/archive/ph_risk/committees/sct/documents/out143_en.pdf

Van Herwijnen, R. (2009). Environmental risk limits for DODMAC and DHTDMACRIVM Letter report 601782029 - https://www.rivm.nl/bibliotheek/rapporten/601782029.pdf

2.3.8. Question 1

Try to look up toxicity literature for the pesticide formulation Roundup and compare the apparent toxicities of the active ingredient glyphosate and the adjuvant surfactant to nontarget organisms. What do you notice?

2.3.8. Question 2

Surfactants do not just accumulate at the air-water interface, but also at water-solvent interfaces. The octanol-water partition coefficient is a critical property for risk assessment, but can you provide reasons why this property is problematic to be derived and applied for surfactants?

2.3.8. Question 3

Although detergents are designed to be readily biodegradable, some compounds appear to be pseudo-persistent as they are still found at detectable levels near the outflow of sewage treatment effluent pipes. What do you think is meant by this?

2.3.9. Food and Feed Additives

In preparation

This page titled 2.3.2.3: Pollutants with specific use is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **2.3: Pollutants with specific use** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

SECTION OVERVIEW

2.3.3: Environmental Chemistry - From Fate to Exposure

2.3.3.1: Environmental compartments

2.3.3.2: Sources of chemicals

2.3.3.3: Pathways and processes determining chemical fate

2.3.3.4: Partitioning and Partitioning Constants

2.3.3.5: Metal Speciation

This page titled 2.3.3: Environmental Chemistry - From Fate to Exposure is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

2.3.3.1: Environmental compartments

3.1.1. Introduction

In order to understand and predict the effects of chemicals in the environment we need to understand the behaviour of chemicals in specific environments and in the environment as a whole. In order to deal with the diversity of natural systems, we consider them to consist of compartments. These are defined as parts of the physical environment that are defined by a spatial boundary that distinguishes them from the rest of the world, for example the atmosphere, soil, surface water and even biota. These examples suggest that three phases: gas, liquid, and solid, are important but compartments may consist of different phases. For example, the atmosphere consists of suspended liquids (e.g., fog) and solids (e.g., dust) as well as gases. Similarly, lakes contain suspended solids and soils contain gaseous and water-filled pore space. In detailed environmental models, each of these phases may also be considered to be a compartment.

The behaviour and fate of chemicals in the environment is determined by the properties of environmental compartments and the physicochemical characteristics of the chemicals. Together these properties determine how chemicals undergo chemical and biological reactions, such as hydrolysis, photolysis and biodegradation, and phase transfer processes such as air-water exchange and sorption.

In this chapter, we first introduce the most important compartments and their most important properties and processes that determine the behaviour of chemical contaminants: the atmosphere, the hydrosphere, sediment, soil, groundwater and biota. The emissions of chemicals into the environment from either point sources or diffuse sources is discussed and the important pathways and processes determining the fate of chemicals. The partitioning approach to phase-transfer processes is presented with sorption as a specific example. The impact of physicochemical properties on partitioning is also discussed.

Other important environmental processes are discussed in sections on metal speciation, processes affecting the bioavailability of metals and organic contaminants and the transformation and degradation of organic chemicals. These sections also include information on the basic methods to measure these processes.

Finally, approaches that are used to model and predict the environmental fate of chemicals, and thus the exposure of organisms to these chemicals are described in section 3.8.

3.1.2. Atmosphere

Authors: Astrid Manders-Groot

Reviewer: Kees van Gestel, John Parsons, Charles Chemel

Leaning objectives:

You should be able to:

- describe the structure of the atmosphere and mention its main components
- describe the processes that determine residence time and transport distances of chemicals in air

Keywords: atmosphere, transport distance, residence time

Composition and vertical structure of atmosphere

The atmosphere of the Earth consists of several layers that have limited interaction. The **troposphere** is the lowermost part of the atmosphere and contains the oxygen that we breathe and most of the water vapor. It contains on average 78% N_2, 20% O_2 and up to 4% water vapor. Greenhouse gases like CO_2 and CH_4 are present at 0.0038 % and 0.0002%, respectively. Air pollutants like ozone and NO_2 have concentrations that are even a factor 1,000-10,000 lower, but are already harmful for the health of humans, animals and vegetation at these concentrations.

The troposphere is 6-8 km high near the poles, about 10 km at mid-latitudes and about 15 km at the equator. It has its own circulation and determines what we experience as weather, with temperature, wind, clouds and precipitation. The lowest part of the troposphere is the **boundary layer**, the part that is closest to the Earth. Its height is determined by the heating of the atmosphere by the Earth surface and the wind conditions and has a daily cycle, determined by the incoming sunlight. It is not a completely

separate layer, but the exchange of air pollutants like O3, NOx, SO2, and xenobiotic chemicals between the boundary layer and the above layers is generally inefficient. Therefore it is also termed the mixing layer.

Above the troposphere there is the **stratosphere**, a layer that is less strongly influenced by the daily solar cycle. It is very dry and has its owns circulation, with some exchange with the troposphere. The stratosphere contains the ozone layer that protects life on Earth against UV radiation and extends to about 50 km altitude. The layers covering the next 50 km are the mesosphere and thermosphere, which are not directly relevant for the transport of the chemicals considered in this book.

Properties of pollutants in the air

Air pollutants include a wide range of chemicals, ranging from metals like lead and mercury to asbestos fibers, polycyclic hydrocarbons (PAH) and chloroform. These pollutants may be emitted into the atmosphere as a gas or as a particle or droplet with sizes of a few nanometer to tens of micrometers. The particles and droplets are termed **aerosol**, or, depending on the measurement method, **particulate matter**. The latter definition is used in air quality regulations. Note that a single aerosol can be composed of several chemical compounds. Once a pollutant is released in the atmosphere, it is transported by diffusion and advection by horizontal and vertical winds and may be ultimately deposited to the Earth's surface by rain (wet deposition), and by sticking to the surface (dry deposition). Large particles may fall down by gravitational settling, a process also called sedimentation. Air pollutants may interact with each other or with other chemicals, particles and water by physical or chemical processes. All these processes will be explained in more detail below. A summary of the relevant interactions is given in Figure 1.

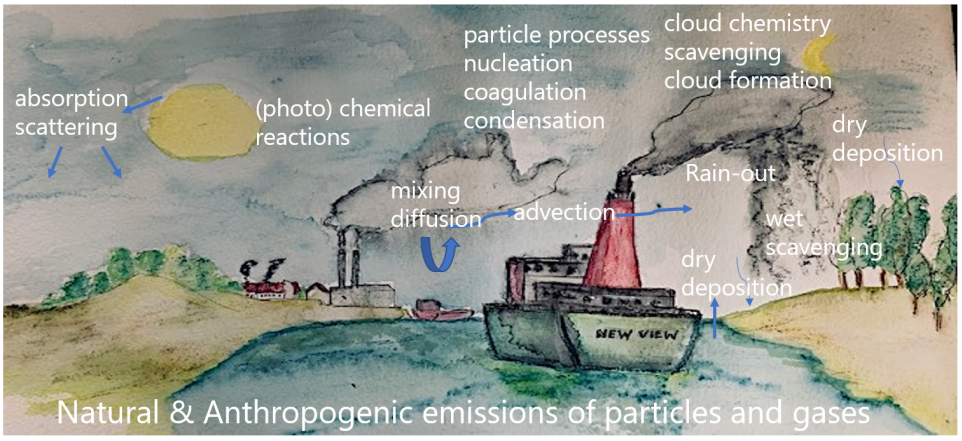


Figure 1. Overview of the most relevant process in the atmosphere related to release and transport of air pollutants (source: Wilma IJzerman).

It is important to realize that air pollutants can have an impact on meteorology itself, by acting as a greenhouse gas, scattering or absorbing incoming light when in aerosol form, or be involved in the formation of clouds. This aspect will not be discussed here.

Meteorology is relevant for all aspects, ranging from mixing and transport to temperature or light dependent reaction rates and absorption of water. Depending on the removal rates, species may be removed with timescales of seconds, like heavy sand particles, to decades or longer, like halogen (Cl, Br)-containing gases, and be transported over ranges of a few meters to crossing the globe several times. Concentrations of gases are often expressed in volume mixing ratios (parts per billion, ppb) whereas for particulate matter the correct unit is in (micro)gram per cubic meter as there is no molecular weight associated to it. For ultrafine particles, concentrations are expressed as numbers of particles per cubic meter, for asbestos, the number of fibers per cubic meter is used.

Physical and chemical processes determining the properties of air pollutants

The properties of air pollutants, like solubility in water, attachment efficiencies to the Earth's surface (water, vegetation, soil) and size of particles, are key elements determining the lifetime and transport distances. These properties may change due to the interaction with other chemicals and with meteorology.

The main physical processes are:

- Condensation or evaporation with decreasing/increasing temperature. A potentially toxic aerosol may thus be covered by semi-volatile chemicals like ammonium sulfate, whilea gas may condensate on an aerosol and be transported further as part of this

aerosol. Some pollutants exist at the same time in aerosol and in gas phase with their partitioning depending on air temperature and relative humidity.
- Gases may cluster to form ultrafine particles of a few nanometers (nucleation) that will grow to larger sizes.
- Particles may grow by coagulation: rapidly moving small particles bump into large slow-moving particles and remain attached to it.
- Particles may take up water (hygroscopicity), leading to a larger diameter.

Chemical conversions include:

- Chemical reactions between gas-phase pollutants and ambient gas, which alters the characteristics of an air pollutant or can lead to the formation of pollutants (e.g. NO_2 being directly emitted by combustion, leading to ozone formation).
- Chemical reactions between aerosols and gases, often involving water attached to aerosol.
- Cloud droplets or water attached to aerosols have their own role in the chemistry of the atmosphere, and gases may diffuse into the water .
- Some pollutants may act as a catalyst.
- Some air pollutants may be degraded by (UV) light (photodegradation).

Pollutants are characterized by their chemical composition but for aerosols also the size distribution of particles is relevant. Note that the conservation of atoms always applies, but particle size distribution and particle number can be changed by physical processes. This has to be kept in mind when concentrations are expressed in particles per volume instead of mass concentrations.

Transport of air pollutants

Several processes determine the mixing and transport of chemicals in the air:

- Diffusion due to the motion of molecules, or Brownian motion of particles (random walk).
- Turbulent diffusion: the mixing due to (small-scale) turbulent eddies which have a random nature, but the strength of this diffusion is related to friction in the flow.
- Advection: the process of transport with the large-scale flow (wind speed and direction).
- Mixing or entrainment of different air masses leads to further mixing of air pollutants over a larger volume. This process is for example relevant when the sun rises, and air in the boundary layer is heated, rises, and mixes with the air in the layer above the boundary layer.

Although the processes of diffusion and transport are well-known, it is not an easy task to solve the equations describing these processes. For stationary point and line sources under idealized conditions, analytical descriptions can be derived in terms of a plume with concentration profile with a Gaussian distribution, but for more realistic descriptions the equations must be solved numerically. For complex flow around a building, computational fluid dynamics is required for an accurate description, for long-range transport a chemistry-transport model must be used.

Wet deposition

Wet deposition comprises the removal processes that involve water:

1. Scavenging of particles by or dissolution of gases in cloud droplets. These cloud droplets may grow to larger size and rain out, thereby removing the dissolved air pollutants from the atmosphere (in-cloud scavenging).
2. Particles below the clouds may be scavenged by falling raindrops (below-cloud scavenging).
3. Occult deposition occurs when clouds are in contact with the surface (mountain areas) and cloud droplets containing air pollutants stick to the surface.

Wet deposition is a very efficient removal mechanism for both small (<0.1 µm diameter) and large aerosols (diameter >1 µm). Aerosols that are hygroscopic can grow in size by absorbing water, or shrink by evaporating water under dry conditions. This affects their deposition rate for wet or dry deposition.

Dry deposition

Dry deposition is partly determined by the gravitational forces on a particle. Heavy particles (≥ 5 µm) fall to the Earth's surface in a process called gravitational settling or sedimentation. In the lowest layer of the atmosphere, air pollutants can be brought close

enough to the surface to stick to it or be taken up. In the turbulent boundary layer, air pollutants are brought close to the surface by the turbulent motion of the atmosphere, up to the very thin laminar layer (*laminar resistance*, only for gases) through which they diffuse to the surface. Aerosols or gases can stick to the Earth's surface or be taken up by vegetation, but they may also rebound. Several pathways take place in parallel or in series, similar to an electric circuit with several resistances in parallel and series. Therefore the **resistance approach** is often used to describe these processes.

Deposition above snow or ice is generally slow, since the atmosphere above it is often stably stratified with little turbulence (high *aerodynamic resistance*), the surface area to deposit on is relatively small (impactors) and aerosols may even rebound to an icy surface (collection efficiency of impactors) to which it is difficult to attach. On the other hand, forests often show high deposition velocities since they induce stronger turbulence in the lowermost atmosphere and have a large leaf surface that may take up gases by the stomata or provide sticking surfaces for aerosols. Deposition velocities thus depend on the type of surface, but also on the season, atmospheric stability (wind speed, cloud coverage) and ability of stomata to take up gases. When the atmosphere is very dry, for example, plants close their stomata and this pathway is temporarily shut down. For particles, the dry deposition velocity is lowest at sizes of 0.1-1 µm.

Re-emission

Once air pollutants are removed from the atmosphere, they can be part of the soil or water compartments which can act as a reservoir. This is in general only taken into account for a limited number of chemicals. Ammonia or persistent organic pollutants may be re-emitted from the soil by evaporation. Dusty material or pollutants attached to dust may be brought back into the atmosphere by the action of wind. This is relevant for bare areas like agricultural lands in wintertime, but also for passing vehicles that bring up the dust on a road by the flow they induce.

Atmospheric fate modelling

Due to the many relevant processes and interactions, the fate of chemical pollutants in the air has to be determined by using models that cover the most important processes. Which processes need to be covered depends on the case study: a good description of a plume of toxic material during an accident, where high concentrations, strong gradients and short timescales are important, requires a different approach than the chronic small release of a factory. Since it would require too heavy numerical simulations to include all aspects, one has to select the relevant processes to be included. Key input for all transport models are emission rates and meteorological input.

When one is interested in concentrations close to a specific source, next to emission rate the effective emission height is important, and processes that determine dispersion: wind speed, atmospheric stability. Chemical reaction rates and deposition velocities should be included when the time horizon is long or when the reactions are fast or deposition velocities are high.

When one is interested in actual concentrations resulting from releases of multiple sources and species over a large area of interest, like for an air quality forecast, the processes of advection, deposition and chemical conversions become more relevant, and input meteorology needs to be known over the area. Sharp gradients close to the individual sources are, however, no longer resolved. In particular rain can be a very efficient removal mechanism, removing most of the aerosol within one hour. Dry deposition is slower, but results in a lifetime of less than a week and transport distances of less than 1,000 km for most aerosols. For some gaseous compounds like halogens and N_2O deposition does hardly play a role and they are chemically inert in the troposphere, leading to very long lifetimes.

To assess the overall long-term fate of a new chemical to be released to the market, the potential concentrations in air, water and soil have to be determined. Ideally, models for air, soil and water are used together in a consistent way, including their interaction For many air pollutants the atmospheric lifetime is short but determines where and in which form they are deposited onto ground and water surfaces, where they may accumulate. This means that even if a concentration in air is relatively low at a certain distance from a source, the deposition of an air pollutant over a year may still be significant. Figure 2 shows an example of annual mean modelled concentrations and annual total deposition of a hypothetical passive (non-reactive) soot-like tracer that is released at 1 kg/hour at a fictitious site in The Netherlands. Annual mean concentrations are small compared to ambient concentrations of particulate matter, but the footprint of the accumulated deposition is larger than that of the mean concentration, since the surface acts as a reservoir. This implies that re-emission to air can be relevant. It may take several years for soil or water before an

equilibrium concentration is reached in these compartments from the deposition input, as different processes and time scales apply. Mountain ranges are visible in the accumulated wet deposition (Alps, Pyrenees), as they are areas with enhanced precipitations.

In addition to spatially explicit modelling, also box models exist that have the advantage that they can make long-term calculations for a continuous release of a species, including interaction between the compartments air, soil and water. They can be used to determine when an equilibrium concentration is reached within a compartment, but these models cannot resolve horizontal concentration gradients within a compartment.

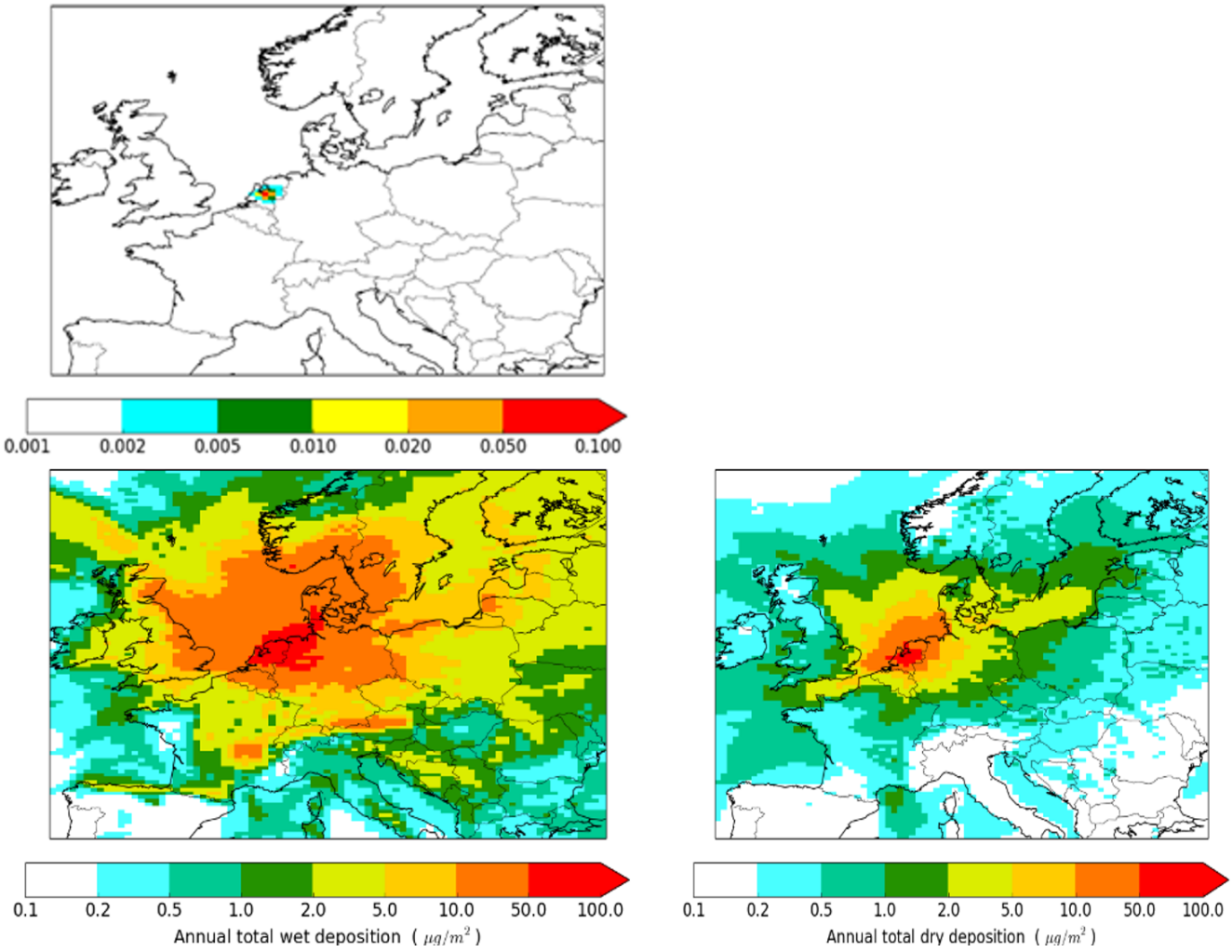


Figure 2. *Constant release of a passive tracer from a point source in The Netherlands. Upper panel shows the annual mean concentration, the lower panel shows the accumulated wet and dry deposition over one year. Note the nonlinear colour scale to cover the large range of values. Source: https://doi.org/10.3390/atmos8050084.*

Learn more

EU air quality, policy and air quality legislation: http://ec.europa.eu/environment/air/index_en.htm

US hazardous air pollutants, including lists of toxics: https://www.epa.gov/haps

Plume dispersion approach: http://courses.washington.edu/cewa567/Plumes.PDF

Chemistry-transport models: www.narsto.org/sites/narsto-dev.ornl.gov/files/Ch71.3MB.pdF

Seinfeld, J., Pandis, S.N. Atmospheric Chemistry and Physics, from air pollution to climate change, Wiley, 2016, covering all aspects.

John, A. C., Küpper, M., Manders-Groot, A. M., Debray, B., Lacome, J. M., Kuhlbusch, T. A. (2017). Emissions and possible environmental implication of engineered nanomaterials (ENMs) in the atmosphere. Atmosphere, 8(5), 84.

3.1.2. Question 1

Which processes determine the concentration of a pollutant in the air very close to its point of emission?

3.1.2. Question 2

When fine particles (<1µm diameter) are released in the lowest 100 m of the atmosphere, how far will they be transported? And what processes do contribute to removal of the particles from the air?

3.1.2. Question 3

When gases are released in the lowest 100 m of the atmosphere, how far do they get?

3.1.3. Hydrosphere

Authors: John Parsons

Reviewers: Steven Droge, Sean Comber

Leaning objectives:

You should be able to:

- describe the most important chemical components and their sources
- describe the most important chemical processes in fresh and marine water.
- be familiar with the processes regulating the pH of surface water.

Keywords: Hydrogen bonding, carbonates, dissolved salts

The properties and importance of water

Water covers 71% of the earth's surface and this water, together with the smaller amounts present as gas in the atmosphere, as groundwater and as ice is referred to collectively as the hydrosphere. The bulk of this water is salt water in the oceans and seas with only a minor part of freshwater being present as lakes and rivers (Figure 1).

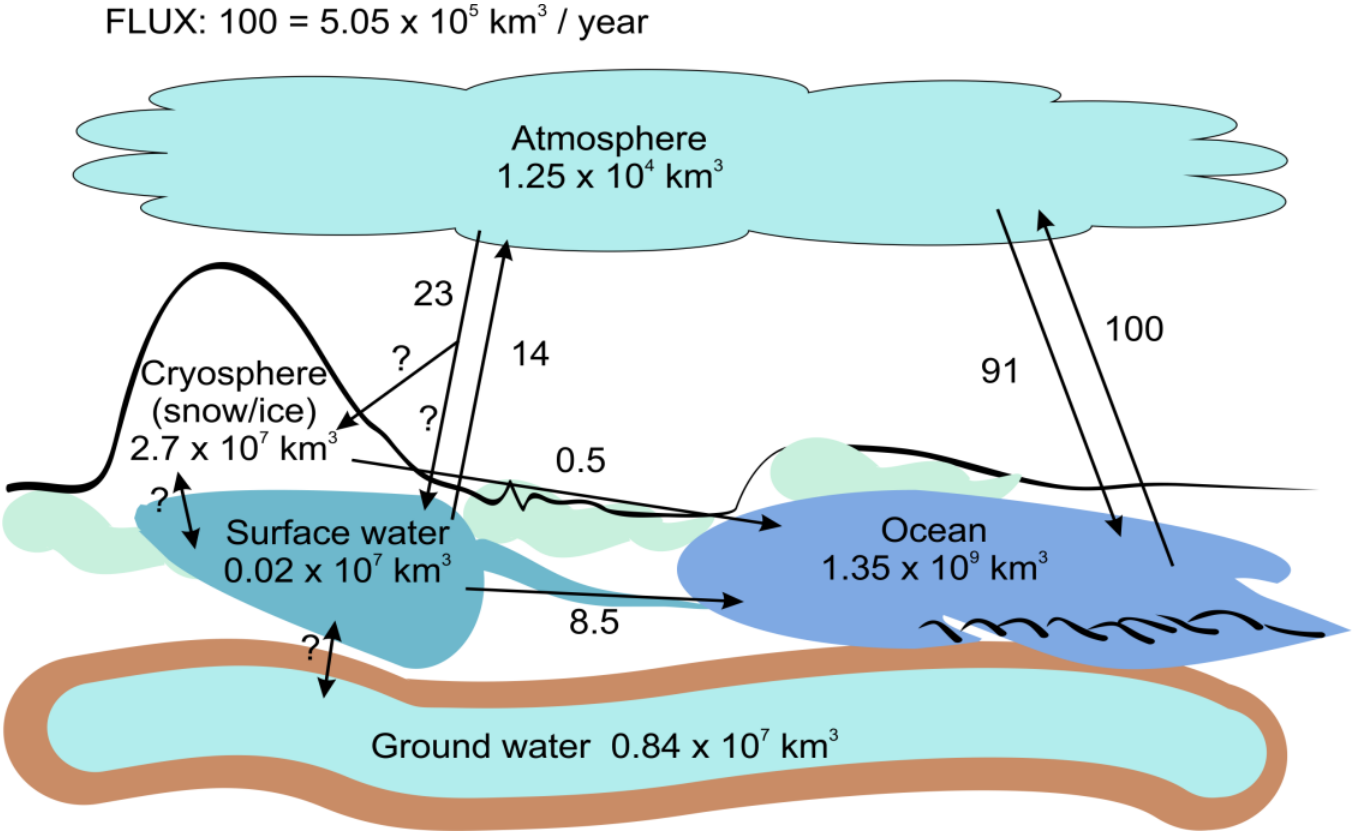


Figure 1. Global hydrological cycle and water balance (arrows are fluxes of water per year). Adapted from Kayane (1992) and Peixoto (1994) by Steven Droge 2019.

Water is essential for life and also plays a key role in many other chemical and physical processes, such as the weathering of minerals and soil formation and in regulating the Earth's climate. These important roles of water derive from its structure as a small but very polar molecule arising from the polarised hydrogen-oxygen bonds (Figure 2). As a consequence, water molecules are strongly attracted by hydrogen bonding, giving it relatively high melting and boiling points, heat capacity, surface tension, etc. The polarity of the water molecule also makes water an excellent solvent for a wide variety of ionic and polar chemicals but a poor solvent for large nonpolar molecules.

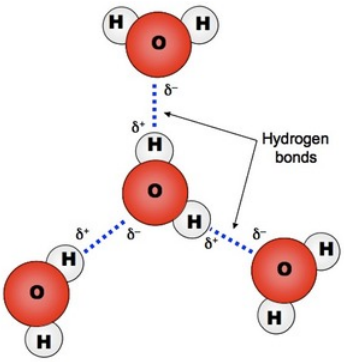


Figure 2. Hydrogen bonding between water molecules

The freshwater environment

As mentioned above, freshwater is only very small proportion of total amount of water on the planet and most of this is present as ice. Since this water is in contact with the atmosphere and the soils and bedrock of the Earth's crust, it dissolves both atmospheric gases such as oxygen and carbon dioxide and salts and organic chemicals from the crust. If we compare the relative compositions of cations in the Earth's crust and the major dissolved species (Table 1) it is clear that these are very different. This difference reflects the importance of the solubility of these components. For ionic chemicals, this depends on both their charge and their size

(expressed as z/r^2, where z is the charge and r the radius of an ion). As well as reflecting the properties of the local crust, the composition of salts is also influenced by precipitation and evaporation and the deposition of sea salt in coastal regions.

Table 1. Comparison of the major cation composition of average upper continental crust and average river water. (*except aluminum and iron from Broecker and Peng (1982))

	Upper continental crust (mg/kg) (Wedepohl 1995*)	**River water (mg/kg) (Berner & Berner 1987*)**
Al	77.4	0.05
Fe	30.9	0.04
Ca	29.4	13.4
Na	25.7	5.2
K	28.6	1.3
Mg	13.5	3.4

The pH of surface water is determined by both the dissolution of carbonate minerals and carbon dioxide from the atmosphere. These components are part of the set of equilibrium reactions known as the carbonate system (Figure 3).

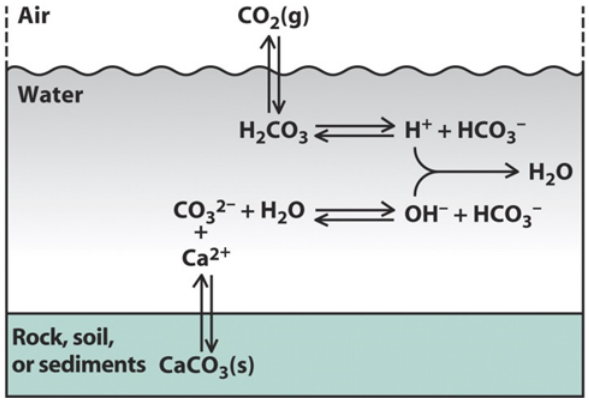


Figure 3. The carbonate system of equilibria regulating the pH of surface water. Source: http://butane.chem.uiuc.edu/pshapley/GenChem1/L26/3.html

At equilibrium with the current atmospheric CO_2 concentration and solid calcium carbonate, the pH of surface water is between 7 and 9 but this may reach more acidic values where soils are calcium carbonate (limestone) poor. This is illustrated by the pH values measured in a river in Northern England, where acidic, organic carbon-rich water at the source is gradually neutralised once the river encounters limestone rich bedrock (Figure 4).

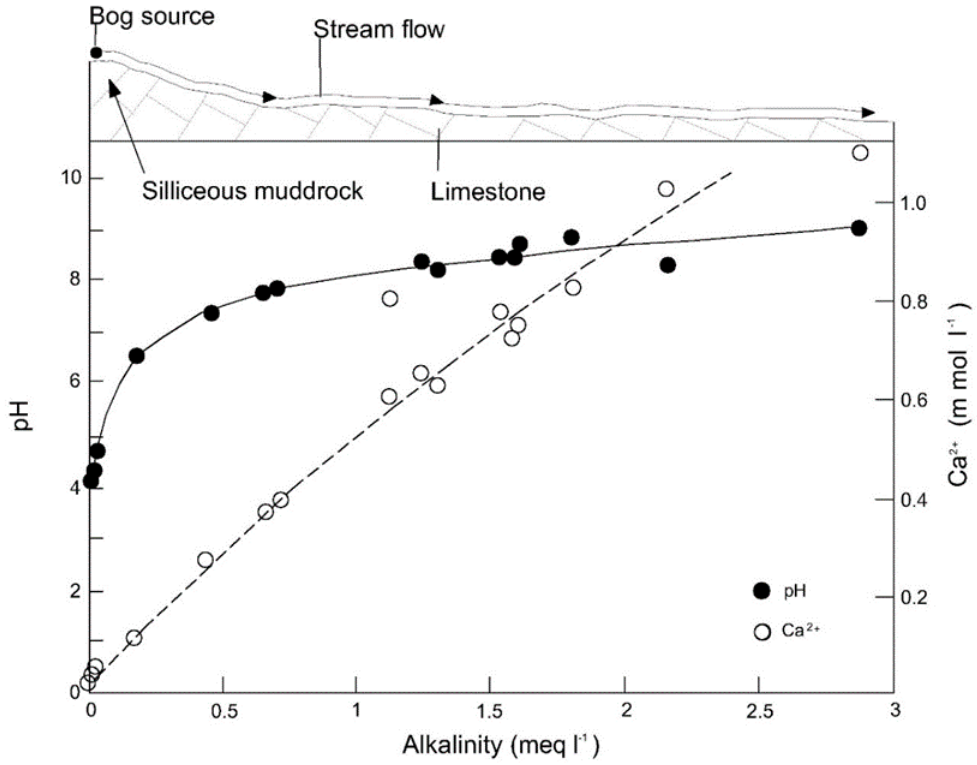


Figure 4. Water chemistry in the Malham Tarn area of northern England, showing the relationship between pH, alkalinity and dissolved calcium as water flows from bog on siliceous mudrock to limestone where the pH is buffered around 8 due to once limestone weathering. Redrawn from Fig. 5.6 in Andrews et al. (2004) by Wilma Ijzerman.

As well as these natural processes, there are human influences on the pH of surface water including acidic precipitation resulting from fossil fuel combustion and acidic effluents from mining activities caused by oxidation and dissolution of mineral sulphides. Regions such as Southern Scandinavia with carbonate-poor soils are particularly vulnerable to acidification due to these influences and this is reflected in for example, reduced fish populations in these vulnerable regions (see Figure 5). More recently, reduced coal burning and the decline in heavy industry is resulting in the recovery of pH values in upland areas across Europe.

https://bio.libretexts.org/@go/page/108672

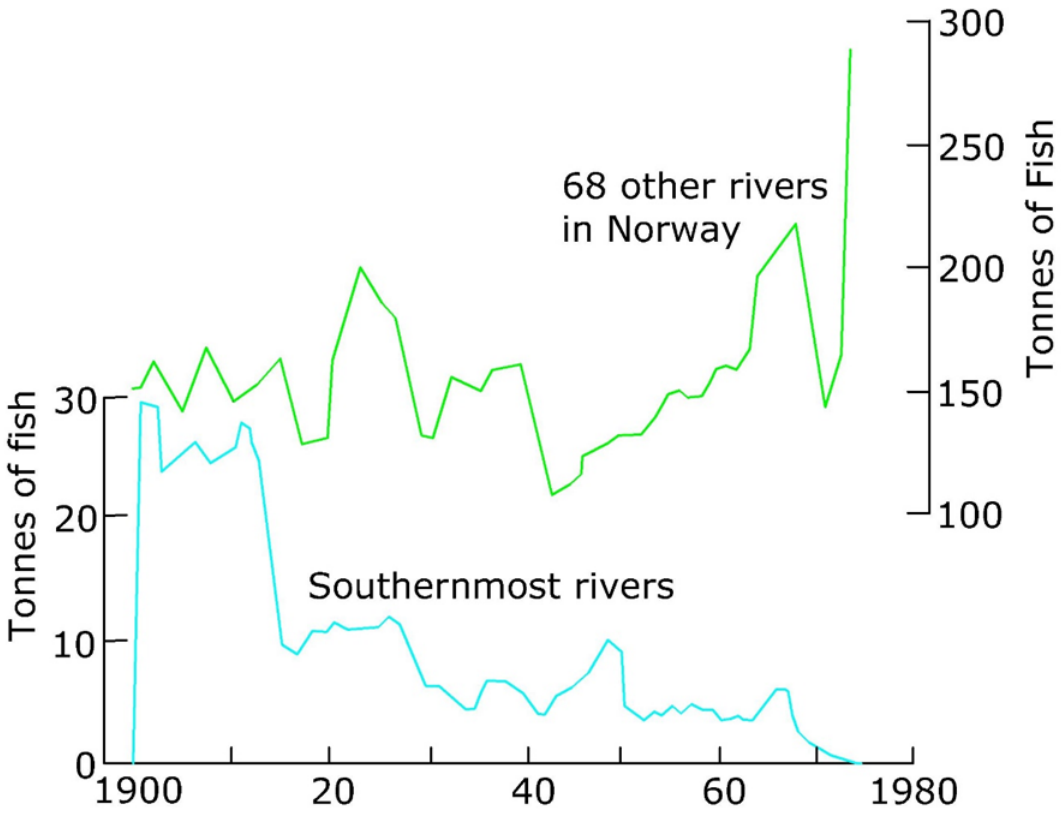


Figure 5. Average catch size of salmon in seven rivers in southern Norway receiving acid precipitation and 68 other rivers that do not receive acid precipitation. Redrawn with data from Henriksen et al. (1989) by Wilma Ijzerman.

Dissolved oxygen is of course essential to aquatic life and concentrations are in general adequate in well mixed water bodies. Oxygen can become limiting in deep lakes where thermal stratification restricts the transport of oxygen to deeper layers, or in water bodies with high rates of organic matter decomposition. This may result in anoxic conditions with significant ecological impacts and on the behaviour of chemical contaminants.

The marine environment

Freshwater eventually moves into seas and oceans where the concentrations of dissolved species are much higher than in the freshwater environment. This is partly due to the effects of evaporation of water from the oceans but is also be due to specific marine sources of some dissolved components. Estuaries are the transition zones where freshwater and seawater mix. These are highly productive environments where increasing salinity has a major impact on the behaviour of many chemicals, for example on the speciation of metals and the aggregation of colloids as a result of cations shielding the negative surface change of colloidal particles (Figure 6). Increasing salinity also affects organic chemicals, with ionic chemicals forming ion pairs, and even reducing the solubility of neutral organics (the so-called salting-out effect). As well as these chemical effects due to increasing salinity, the lowering of flow rates in estuaries leads to the deposition of suspended particles.

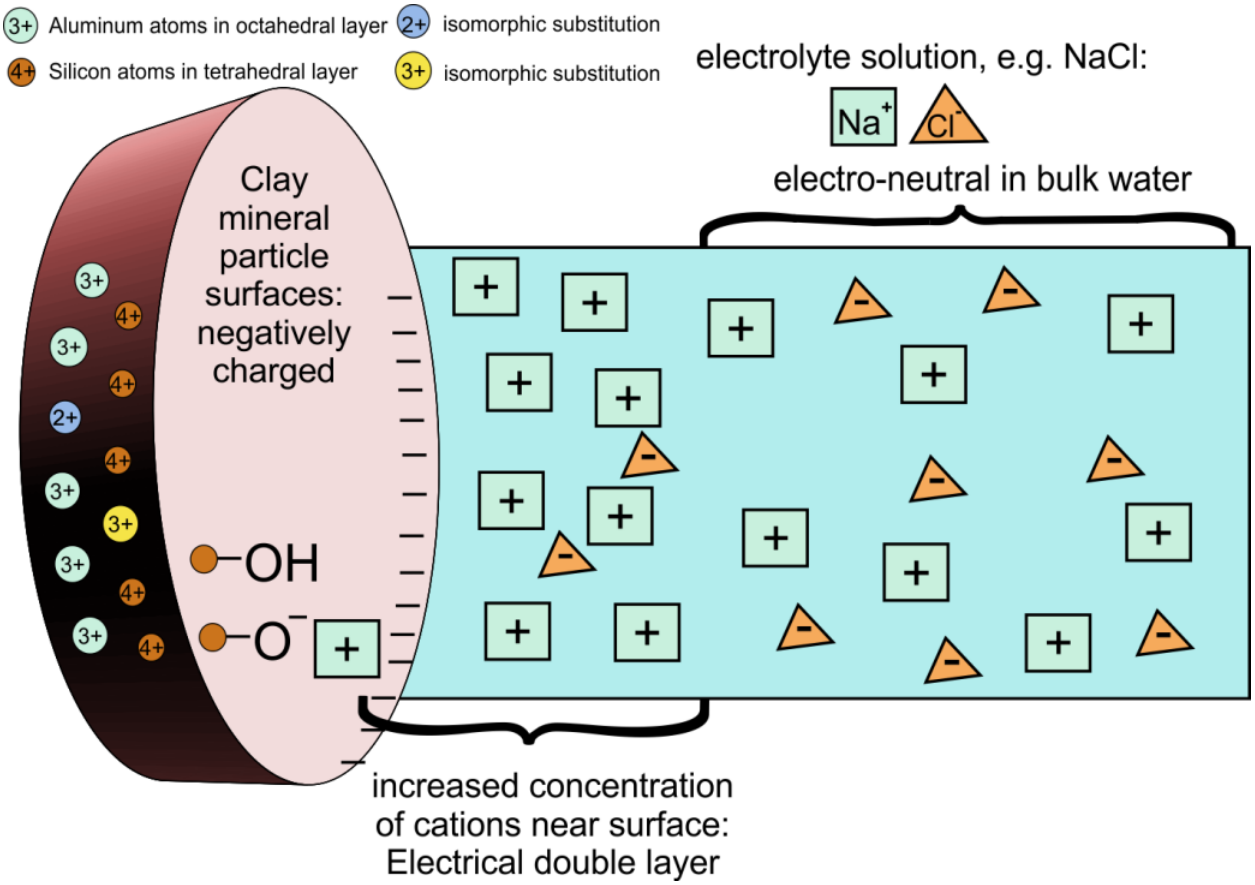


Figure 6. The Electrical Double Layer (EDL), comprising a fixed layer of negative charge on a clay particle (due to isomorphic substitutions and surface acids) and a mobile ionic layer in solution. The latter is caused because positive ions are attracted to the particle surface. Note that with increasing distance from the particle surface the solution approaches electrical neutrality. (Source Steven Droge 2019)

Since the concentrations of pollutants are in general lower in the marine environment than in the freshwater environment, concentrations in estuaries decrease as freshwater is diluted with seawater. Measuring salinity at different locations in estuaries is a convenient way to determine the extent of this dilution. Components that are present in higher concentrations in seawater will of course show an increase with salinity unless. Plotting salinity against the concentrations of chemicals at different locations can yield information on whether they behave conservatively (i.e. only undergoing mixing) or are removed by processes such as degradation or partitioning into the atmosphere or sediments. Figure 7 shows examples of plots expected for conservative chemicals and those that are either removed in the estuary or have local sources there. Models describing the behaviour of chemicals in estuaries can be used with these data to derive the rates of removal or addition of the chemical in the system.

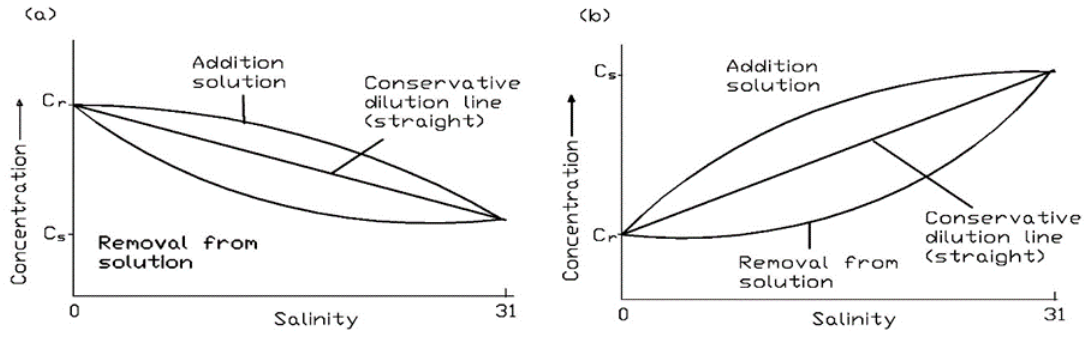


Figure 6.3 in Andrews et al. (2004) by Wilma Ijzerman.

The open ocean is sufficiently mixed for the composition of major dissolved constituents to be fairly constant, except in local situations as a result of upwelling of deep nutrient-rich waters or the biological uptake of nutrients. In coastal regions the

concentrations of chemicals and other components originating from terrestrial sources may also be locally higher. The major components in seawater are listed in Table 2 with their typical concentrations.

Table 2. Major ion composition of freshwater and seawater.

	Seawater (mmol/L) (Broecker and Peng, 1982)	**River water (mmol/L) (Berner and Berner, 1987)**
Na^+	470	0.23
Mg^{2+}	53	0.14
K^+	10	0.03
Ca^{2+}	10	0.33
HCO_3^-	2	0.85
SO_4^{2-}	28	0.09
Cl^-	550	0.16
Si	0.1	0.16

These concentrations may be higher in waterbodies that are partly or wholly isolated from the oceans and are impacted by evaporative losses of water (e.g. Mediterranean, Baltic, Black Sea). In extreme case, concentrations of salts may exceed their solubility product, resulting in precipitation of salts in evaporate deposits.

As is the case in freshwater, carbonates play an important role in regulating the ocean pH. The fact that the oceans are supersaturated in calcium carbonate makes it possible for a variety of organisms to have calcium carbonate shells and other structures. The important processes and equilibria involved are illustrated in Figure 8. There is concern that one of the most important effects of increasing atmospheric carbon dioxide will be lowering of ocean pH to values that will result in destabilisation of these carbonate structures.

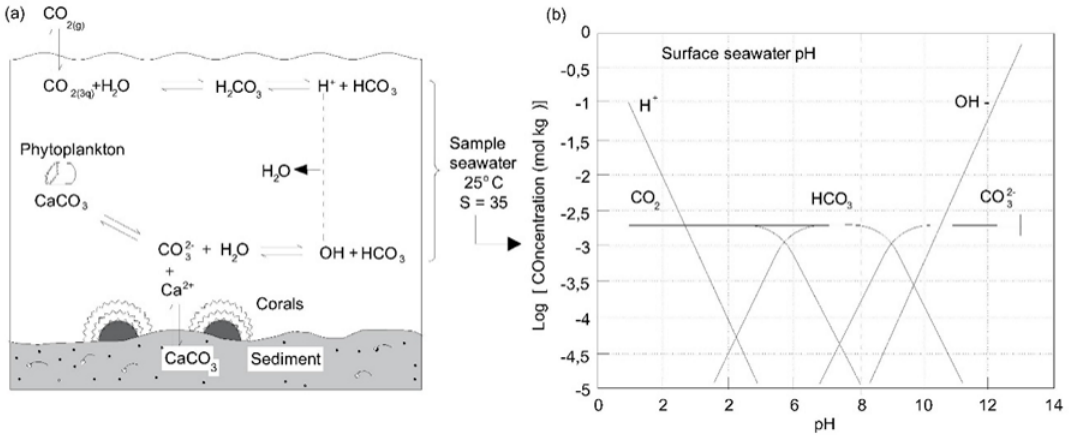


Figure 6.8 in Andrews et al. (2004) by Wilma Ijzerman.

References

Andrews, J.E., Brimblecombe, P., Jicketts, T.D., Liss, P.S., Reid, B.J. (2004). An Introduction To Environmental Chemistry, Blackwell Publishers, ISBN 0-632-05905-2.

Baird, C., Cann, M. (2012). Environmental Chemistry, Fifth Edition, W.H. Freeman and Company, ISBN 978-1429277044.

Berner, E.K., Berner, R.A. (1987). Global water cycle: geochemistry and environment, Prentice-Hall.

Broecker, W.S., Peng, T.S. (1982). Tracers in the Sea, Lamont-Doherty Geol. Obs. Publ.

Henriksen, A., Lien, L., Rosseland, B.O., Traaen, T.S., Sevaldrud, I.S. (1989). Lake Acidification in Norway: Present and Predicted Fish Status. Ambio 18, 314-321

Wedepohl, K.H. (1995). The composition of the continental crust, Geochimica Cosmochimica Acta 59, 1217-1232.

3.1.3. Question 1

The concentrations of dissolved salts in rivers and lakes is determined by three processes. Describe these processes and the characteristic composition of waterbodies where one of these processes is dominant.

3.1.3. Question 2

In addition to the well-known effects on global climate, increasing atmospheric CO_2 is also expected to have an impact on the pH of the oceans. Which processes are responsible for determining the pH of the oceans?

3.1.3. Question 3

How could increasing atmospheric CO_2 affect these processes?

3.1.3. Question 4

What effects could increasing acidity of the oceans have on marine organisms?

3.1.4. Sediment

In preparation

3.1.5. Soil

Author: Kees van Gestel

Reviewers: John Parsons, Jose Alvarez Rogel

Learning goals

You should be able to

- describe the main components of which soils consist
- describe how soil composition influences properties that may affect the fate of chemicals in soil

Keywords:

Particle size distribution, Porosity, Minerals, Organic matter, Cation Exchange Capacity, Water Holding Capacity

Introduction

Soil is the upper layer of the terrestrial environment that serves as a habitat for organisms and medium for plant growth. In addition, it also plays an important role in water storage and purification and helps to regulate the Earth's atmosphere (e.g. carbon storage, gas fluxes, …).

Soils are composed of three phases (Figure 1).

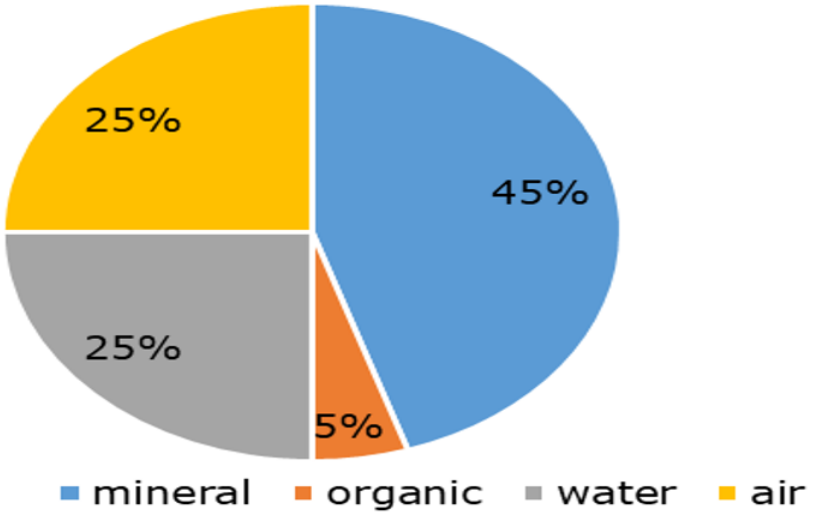


Figure 1. Average composition of soil (in volume %).

The *solid phase* is formed by mineral and organic components. Mineral components appear in different particle sizes from coarse particles (sand), intermediate (silt) and fine (clay) which combination determine soil texture. The particles can be arranged to form porous aggregates; soil pores being filled with air and/or water. The proportion of air in soils depends on soil moisture content. The composition of the soil solid phase may be quite variable.

The *gaseous phase* has a similar composition as the air, but due to the respiration of plant roots and the metabolic activity of soil microorganisms, O_2 content generally is lower and CO_2 content higher. Exchange of gases between soil pores and atmospheric air takes place by diffusion. Diffusion proceeds faster in dry soil and much slower when soil pores are filled with water.

The *liquid phase* of the soil, the *soil solution* or *pore water*, is an aqueous solution containing ions (mainly Na^+, K^+, Ca^{2+}, Cl^-, NO_3^-, SO_4^{2-}, HCO_3^-) from dissolution of a variety of salts, and also contains dissolved organic carbon (DOC, also referred to as dissolved organic matter, DOM). The soil solution is part of the hydrological cycle, which involves input from among others rain and irrigation, and output by water uptake by plants, evaporation, and drainage to ground and surface water. The soil solution acts as a carrier for the transport of chemicals in soil, both to plant roots, soil microorganisms and soil animals and to ground and surface water.

Soil solids

The soil *solid phase* consists of mineral and organic soil particles. Based on their size, the mineral particles are divided into sand (63-2000 µm), silt (2- 63 µm), and clay (<2 µm). With increasing particle size, the specific surface area decreases, pore size increases and water retention capacity decreases. The *sand* fraction mainly consists of quartz (SiO_2) and does not have any sorption properties because the quartz crystals are electrically neutral. Sandy soils have large pores, so a low capacity to retain water. In soils with a high *silt* fraction, smaller pores are better represented, giving these soils a higher water retention capacity. Also the silt fraction has no adsorptive properties. *Clays* are aluminium silicates, lattices composed of SiO_4 tetrahedrons and $Al(OH)_6$ octahedrons. Upon the formation of clay particles, *isomorphic substitution* occurred, a process in which Si^{4+} was replaced by Al^{3+}, and Al^{3+} by Mg^{2+}. Although having similar diameters, these elements have different valences. As a consequence, clay particles have a negative charge, making positive ions to accumulate on their surface. This includes ions important for plant growth, like NH_4^+, K^+, Na^+ and Mg^{2+}, but also cationic metals (Figure 2). Many other minerals have pH-dependent charges (either positive or negative) which are also important in binding cations and anions.

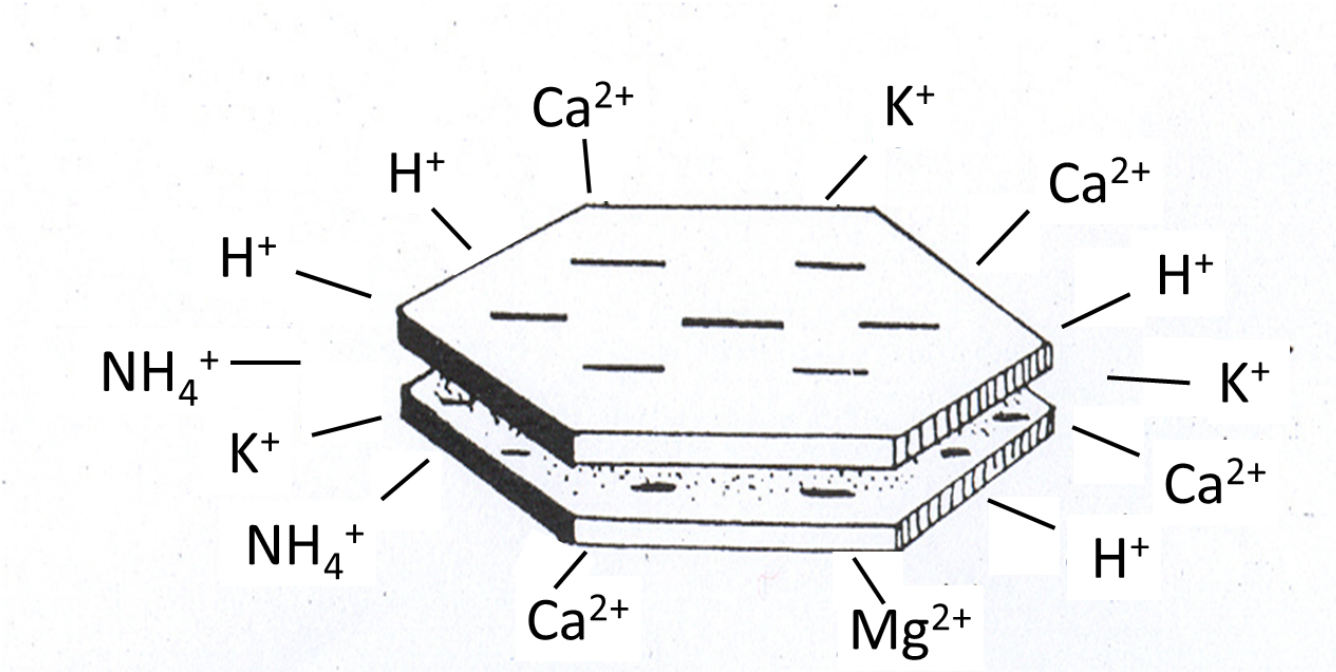


Figure 2. Schematic representation of a clay particle. Due to its negative charge, cations will accumulate near the surface of the clay particle.

In addition to mineral particles, soils also contain organic matter, which includes all dead plant and animals remains and their degradation products. Living biota is not included in the soil organic matter faction. Organic matter is often divided into: 1. *humin*, non-dissolved organic matter associated with clay and silt particles, 2. *humic acids* having a high degree of polymerization, and 3. *fulvic acids* containing more phenolic and carboxylic acid groups. Humic and fulvic acids are water soluble but their solubility depends on pH. For example, humic acids are soluble at alkaline pH but not at acidic pH. The dissociation of the phenolic and carboxylic groups gives the organic matter also a negative charge (Figure 3), the density of which increases with increasing soil pH. The soil organic matter acts as a reservoir of nitrogen and other elements, provides adsorption site for cations and organic chemicals, and supports the building of soil aggregates and the development of soil structure.

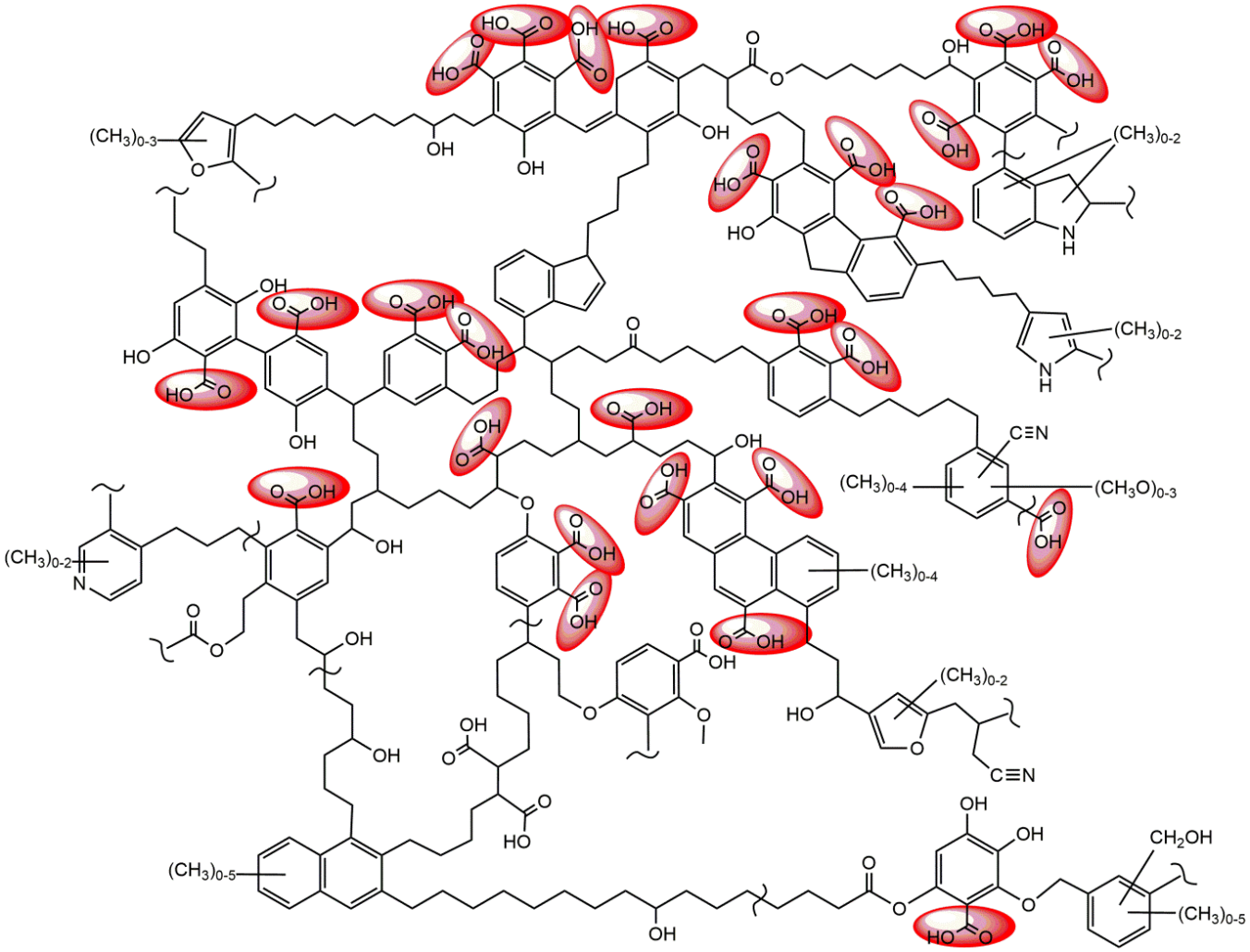


Figure 3. Proposed structure of a humic acid molecule. The phenolic and carboxylic acid groups on the molecule may dissociate depending on the pH of the soil, giving rise to negative sites on the molecule. With this, humic acid contributes to the Cation Exchange Capacity of the soil. Adapted from Schulten & Schnitzer (1997) by Steven Droge.

The binding of cations to the negatively charged sites on the soil particles is an exchange process. The degree of cation accumulation near soil particles depends on their charge density, the affinity of the cations to the charged surfaces (which is higher for bivalent than for monovalent cations), the concentration of ions in solution (the higher the concentration of a cation in solution, the higher attraction to soil particles), etc. Due to their binding to charged soil particles, cations are less available for leaching and for uptake by organisms. The *Cation Exchange Capacity* (CEC) is commonly used as a measure of the number of sites available for the sorption of cations. CEC is usually expressed as $cmol_c/kg$ dry soil. Soils with higher CEC have a higher capacity to bind cations, so cationic metals show a lower (bio)availability in high CEC soils (see the Section on metal speciation). CEC depends on the content and type of clay minerals, with montmorillonite having a higher CEC than e.g. kaolinite, organic matter content and pH of the soil. In addition to clay and organic matter, also aluminium and iron oxides and hydroxides may contribute to the binding of cations to the soil.

Soil water

The transport of water through soil pores is controlled by gravity, and by suction gradients which are the result of water retention by capillary and osmotic processes. Capillary binding of water is stronger in smaller soil pores, which explains why clayey soils have higher water retention capacities than sandy soils. The osmotic binding of water increases with increasing ionic strength, and is especially high close to charged soil particles like clay and organic matter where ions tend to accumulate.

The stronger water is retained by soil, the lower its availability is for plants and other organisms. The strength by which water is retained depends on moisture content, because 1. at decreasing moisture content the ionic strength of the soil solution and therefore osmotic binding increases, 2. when soil moisture content decreases the larger soil pores will be emptied first, leading to increasing capillary retention of the remaining water in smaller pores. Water retention curves describe the strength with which water is retained as a function of total water content and in dependence of the composition of the soil. Figure 4 shows pF curves for three different soil types.

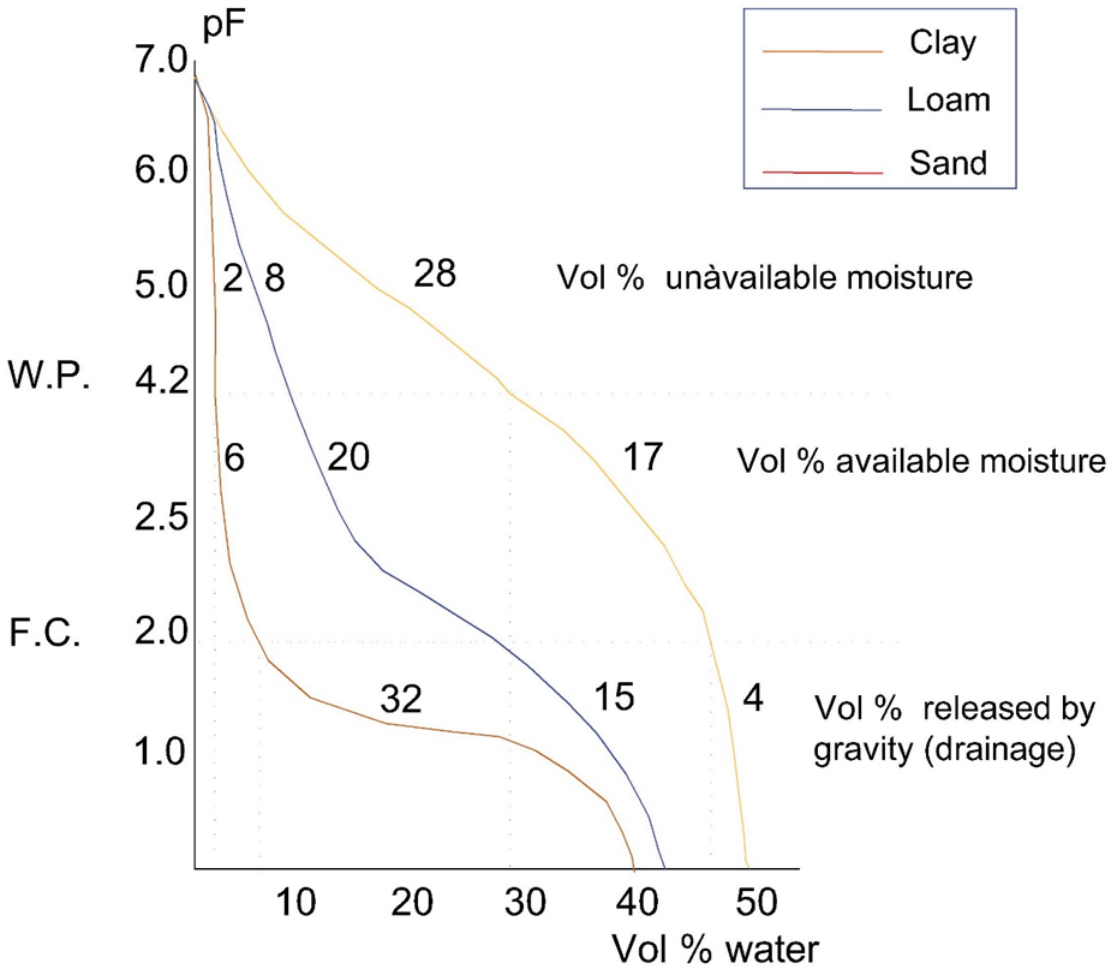


Figure 4. pF curves showing the retention of water by three different soil types. pF is the log of the force with which water is retained, expressed in hPa. W.P. = wilting point, F.C. = field capacity. Source: Wilma IJzerman.

A pF value of 2.2-2.5 corresponds with a binding strength of 200 to 300 hPa. This is called *field capacity*; water is readily available for plant uptake. At pF 4.2 (15,000 hPa), water is strongly bound in the soil and no longer available for plant uptake; this is called the *wilting point*. For soil organisms, not the total water content of a soil is of importance but rather the content of available water. Water retention curves may be important to describe the availability of water in soil. Toxicity tests with soil organisms are typically performed at 40-60% of the *water holding capacity* (WHC) of the soil, which corresponds with field capacity.

References/further reading

Schulten, H.-R., Schnitzer, M. (1997). Chemical model structure for soil organic matter and soils. Soil Science 162, 115-130.

Blume, H.-P., Brümmer, G.W., Fleige, H., Horn, R., Kandeler, E., Kögel-Knabner, I., Kretzschmar, R., Stahr, K., Wilke, B.-M. (2016). Scheffer/Schachtschabel Soil Science, Springer, ISBN 978-3-642-30941-0

3.1.5. Question 1

What are the main components of the solid phase of a soil?

3.1.5. Question 2

When considering two different soils, a sandy and a clayey soil, which one has the highest available volume of water when both soils have the same total water content? Explain your answer.

3.1.5. Question 3

What is the agricultural and environmental importance of the cation exchange capacity of soils?

3.1.5. Question 4

Explain the role of clay minerals and organic matter in the retention of cationic metals in soils.

3.1.5. Question 5

Which two factors explain the fact that the availability of cationic metals increases with decreasing soil pH?

3.1.6. Groundwater

(draft)

Author: Thilo Behrends

Reviewer: Steven Droge, John Parsons

Leaning objectives:

You should be able to:

- understand the significance of redox reactions for the fate of potentially toxic compounds in groundwaters and aquifers.
- apply the Nernst equation to assess the feasibility of redox reactions.

Keywords: Aquifer, Nernst equation, electron transfer, redox potential, half reactions

Introduction

Some definitions conceive all water beneath the earth's surface as groundwater while others restrict the definition to water in the saturated zone. In the saturated zone the pores are completely filled with water in contrast to the undersaturated zone in which some pores are filled with gas and capillary action are important for moving water. Geological formations, which host groundwater in the saturated zone, can be classified as 'aquifer', ' aquitard', or 'aquifuge' depending on their permeability. In contrast to aquitard and aquifuge, which have a low permeability, an aquifer permits water to move in significant rates under ordinary field conditions. Aquifers typically have a high porosity and the pores are well connected with each other. Examples of aquifers include sedimentary layers of sand or gravel, carbonate rocks, sandstones, volcanic rocks and fractured igneous rocks. The redox chemistry discussed in this chapter is focusing on aquifers in sedimentary formations.

Groundwaters are an important source for drinking water and the quality of groundwater is, therefore, of high importance for protecting human health. However, aquifers also represent a habitat for bacteria and aquatic invertebrates and are, therefore, also an object for ecotoxicological studies. Furthermore, groundwater can act as a transportation pathway connecting different environmental compartments e.g. soils with rivers or oceans. Groundwater thus plays a role in the distribution of contaminants in the environment.

Transport of contaminants in aquifers

The movement of a chemical in groundwater is controlled by three processes: advection, dispersion and reaction. Advection is the transport of a chemical in dissolved form together with the groundwater flow. When a chemical is released from a point source into groundwater with a constant flow direction, a plume is forming downstream of the source. The spreading of the chemical is due to dispersion. There are two reasons for this spreading: First, molecular diffusion causes transport of the chemical independently from advection; Second, differences in groundwater velocities at different scales causes mixing of the groundwater (mechanical dispersion) in the direction of groundwater flow but also perpendicular to it. Several process can retard the transport of chemicals or can cause its removal from the system (e.g. degradation). For the mobility of a chemical, the distribution between immobile solid phase and moving liquid phase is of key importance in groundwater (see chapter 3.4). There are several processes which can lead to the degradation of a compound in aquifers. Microbial activity can contribute to the degradation of chemicals but also abiotic reactions can be of importance. For some chemical, redox reactions are relevant which are discussed in the following section.

Redox reactions in aquifers

Many elements are redox-sensitive under environmental conditions. This means they occur naturally in different 'redox states'. For example oxidation or reduction of carbon plays a pivotal role in the energy metabolism of living organisms and carbon occurs in oxidation states from +IV in CO_2 (because the two oxygen atoms both count as -II ((because oxygen is more electronegative than carbon)), and the total molecule should balance out) to -IV in CH_4 (because each H-atom counts as +I ((because hydrogen is less electronegative than carbon))). Also potentially toxic elements, such as arsenic, are found in nature in different oxidation states. Important oxidation states of arsenic are +V, (e.g. AsO_4^{3-}, arsenate), + III (e.g. AsO_3^{-3}, arsenite), 0 (elemental arsenic or arsenic associated with sulfide as in FeAsS, arsenopyrite). Arsenic can also have negative oxidation states when it forms arsenides such as $FeAs_2$ (löllingite). Bioavailability, toxicity and mobility of redox sensitive elements are usually strongly dependent on their oxidation state. For example, arsenite tends to be more toxic and more mobile than arsenate. For this reason, assessing the redox state of potentially toxic elements is an important element of environmental risk assessment of groundwater.

Organic contaminants can also undergo redox transformations. At the earth surface, when oxygen is present, (photo-)oxidation is an important degradation pathway for organic contaminants. In subsurface environments, when oxygen concentrations are often very low (anoxic conditions), reduction can play an important role in degradation pathways. For example, the reductive dehalogenation of chlorinated hydrocarbons or the reduction of nitroaromatic compounds have been extensively investigated. The reduction of these compounds can be mediated by microorganisms but they can also occur abiotically on solid surfaces present in the subsurface. In any case, reduction of organic contaminants is only possible when the reaction is thermodynamically feasible. For this reason it is necessary to know the redox conditions in, for example, an aquifer.

Quantitative assessment of redox reactions

As the name indicates, redox reactions combine oxidation of one constituent in the system with the reduction of another and, hence, involve electron transfer. The oxidation of arsenite with elemental oxygen to arsenate has following stoichiometry:

$$AsO_3^{3-} + 0.5\,O_2 \leftrightarrow AsO_4^{4-}$$

It is important that the stoichiometries of redox reactions are not only charge- and mass-balanced but also electron-balanced. Here, arsenic releases two electrons when going from oxidation state +III to +V (arsenite becomes oxidized to arsenate) while one oxygen atom takes up the two electrons and goes from oxidation state 0 to -II (elemental oxygen becomes reduced). For this reaction an equilibrium constant can be obtained and based on the activities (or concentrations) of dissolved reactants and products it can be evaluated whether the reaction is in equilibrium or in which direction the reaction is thermodynamically favorable.

When a natural system contains several different redox-active constituents, a large number of possible redox reactions can be formulated and evaluated separately. In this situation it is more convenient to formulate and compare *half reactions*. For examples, the oxidation of arsenite with oxygen can be split up into the reactions of arsenic and oxygen.

$$H_3SO4 + 2\,e^- + 2\,H^+ \leftrightarrow H_3SO_3 + H_2O \; E_h^\circ = 0.575\,V$$

$$O_2 + 4\,e^- + 4\,H^+ \leftrightarrow 2\,H_2O \; E_h^\circ = 1.229\,V$$

Half reactions are typically formulated as reduction reactions (electrons are on the left hand side of the reaction). The E_h^o is the standard redox potential and represents the electrical potential, which would be measured in a standardized electrochemical cell

which contains on one side, H_3AsO_4, H_3AsO_3 and H^+, all with activities of 1 mol l^{-1}, and a solution containing 1 mol l^{-1} H^+ in equilibrium with H_2 gas with a pressure of 1 bar, on the other side.

In natural environments the pH is usually not 0 and the activities of arsenite and arsenate are not 1 mol l^{-1}. The redox potential, E_h under these conditions can be calculated using the Nernst equation:

$$E_h = E_h^\circ + \frac{RT}{zF} ln \frac{ox}{red} = 0.575\ V + \frac{8.314\ J\ mol^{-1}\ K^{-1}\ 298\ K}{2 * 96485\ mol\ C^{-1}} ln \frac{\{H_3AsO_4\}\{H^+\}^2}{\{H_3AsO_3\}}$$

where:

R is the ideal gas constant (8.314 J mol^{-1} K^{-1}),

T the temperature in K,

z is the number of electrons which are transferred in the reaction,

F the Faraday constant (96485 mol C^{-1}).

In the ratio ox/red, 'ox' represents the activities or pressures of the constituents on the right hand side of the half reaction, whereby the stoichiometric factor becomes the corresponding exponent, while 'red' represents the right hand side of the half reaction.

The redox potentials of different half reactions can be compared:

- The half reaction with the higher redox potential provides the electron acceptor in the thermodynamically favorable redox reaction,
- The half reaction with the lower potential provides the electron donor.

In other words, it is thermodynamically favorable that the half reaction with the high potential proceeds from left to right and the half reaction with the low potential from right to left.

Redox conditions in aquifers

The redox conditions in an aquifer depends on the inherited inventory of oxidants and reductants during the formation of the geological formation and the processes which have been occurring throughout its history. Oxidants and reductants can have entered the aquifer by diffusion or with the infiltrating water and slowly progressing redox reaction can have modified the assemblage of oxidants and reductants. In the absence of (microbial) catalysis redox reactions often have very slow kinetics. Furthermore, due to photosynthesis, redox reactions are not in equilibrium at the earth's surface and the upper part of the underlying subsurface. As a consequence, the redox conditions in an aquifer can usually not be represented in one unique redox potential. This implies that values obtained for groundwaters with electrochemical measurements, e.g. potentiometric measurements using redox electrodes, might be not representative for the redox conditions in the aquifer. Furthermore, relevant half reactions in the aquifer often involve solids (heterogeneous reactions) with low solubility, implying that the concentrations in solution (for example of Fe^{3+}) are too low to be detected. Hence, evaluating the redox conditions in subsurface environments is often challenging.

Oxygen concentrations in groundwaters are often virtually zero, as oxygen in infiltrating rain water or entering the subsurface by molecular diffusion is often consumed before it can reach the aquifer. Hence, 'reducing conditions' typically prevail in aquifers. The redox potential measured in a system may reflect the dominant electron acceptors besides oxygen that are present in the system (Figure 1).

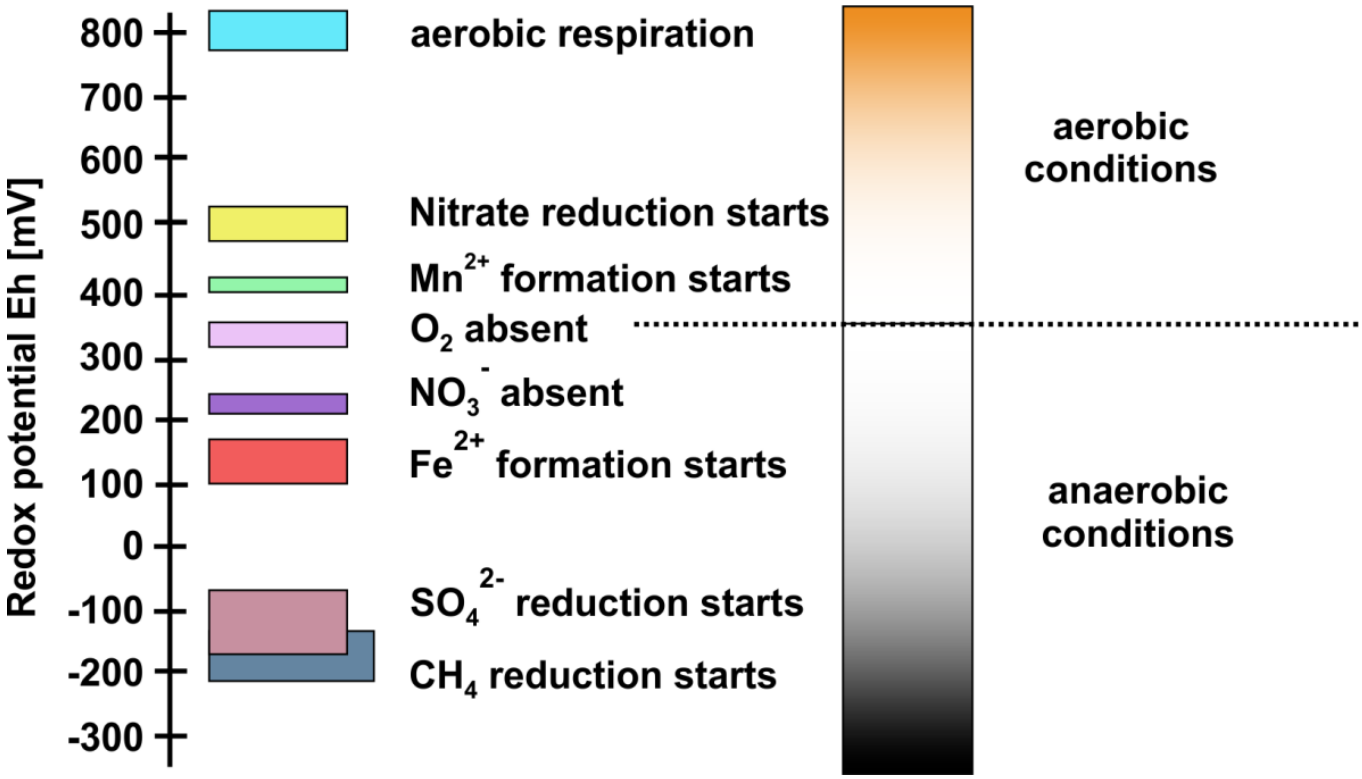


Figure 1. Redox potential scale, ranging from oxic to anoxic conditions. Even when oxygen is absent, the concentrations of available alternative oxidant materials (NO^{3-} and Fe^{2+}) may render a positive redox potential. Methanogenesis (CH_4 formation) only initiates at very low negative redox potentials. (Source: Steven Droge)

In sediments or sedimentary rocks, redox reactions after deposition are predominately driven by the oxidation of organic matter, which entered the sediment during its deposition. However, the aquifer might also have received dissolved organic matter via infiltrating water. The oxidation of organic matter is predominately microbially mediated and predominately coupled to the reduction of elemental oxygen (if present). However, when elemental oxygen is depleted, which is usually the case, other electron acceptors are used by microorganisms. Relevant electron acceptors (see Figure 1) in anoxic environments include:

- Nitrate (in dissolved form),
- Mn(IV) (as solid surface),
- Mn(III) (as solid surface),
- Fe(III) (as solid surface),
- Sulphate (in dissolved form).

Nitrate and sulphate can be present in dissolved form while Mn(IV), Mn(III), Fe(III) occur as solids with low solubility. The (hydr)oxide solids of these metals, such as goethite (FeOOH) or manganite (MnOOH) are mostly accessible for microbial reduction while Mn(III) or Fe(III) in silicates can only be partially reduced or are not bioavailable for reduction. When also these electron acceptors run short, methanogenesis can be initiated.

Microorganisms, which reduce sulphate, Mn or Fe(III), can use the products of fermentative organisms. These fermentative organisms produce short-chain fatty acids, such as acetate or lactate, but often also release hydrogen gas. That is, hydrogen concentrations in groundwater reflect a steady state of hydrogen production and consumption, and are typically limited by the rates of production. As a consequence, hydrogen concentrations in groundwater are often at the physiological limit of the consuming organism. The concentrations are just sufficient to allow the organism to conserve energy from oxidizing the hydrogen. This limit increases according to the sequence of electron acceptors (Figure 1): nitrate reduction < Mn reduction < Fe reduction < sulphate reduction < methanogenesis when the corresponding compounds are present in relevant amounts or concentrations. For this reason, concentrations of dissolved hydrogen can be a useful indicator to identify the dominant, anaerobic respiration pathway in an aquifer. For example, one can determine whether sulphate reduction is enabled or methanogenesis has set in. The hydrogen

concentrations in the groundwater can also directly be used to assess whether the microbial reduction of metals, metalloids, chlorinated hydrocarbons, nitro aromatic compounds or other organic contaminants is feasible.

The reduction of Fe(III)(hydr)oxides are sulphate leads to the formation of Fe(II) and sulphide, which, in turn, typically results in the precipitation of ferrous solids such as $FeCO_3$ (siderite), FeS (mackinawite) or FeS_2 (pyrite). These Fe(II) containing minerals often play an important role in the abiotic reduction of organic or inorganic contaminants in aquifers. When the composition of the groundwater and the mineral assemblage is known, the Nernst equation can be used to calculate the redox potential of relevant half reactions in the aquifer. These redox potential can be then used for evaluating whether reduction of potentially toxic compounds is possible or not. For example, the half reaction for the reduction of an amorphous ferric iron hydroxide coupled to the precipitation of siderite is given by:



$$Fe(OH)_3(s) + e^- + H^+ + H_2CO_3 \leftrightarrow FeCO_3(s) + H_2O \; E_h^\circ = 0.579\,V$$

At given pH and carbonic acid concentration, the corresponding redox potential can be calculated using the Nernst equation. This redox potential can be compared to that obtained from the Nernst equation for the reductive dichlorination of tetrachloroethylene ($Cl_2C{=}CCl_2$)



$Cl_2C = CCl_2 + 2e^- + H^+ \leftrightarrow Cl_2C = CHCl + Cl^- \; E_h^\circ = 0.70\,V$

With this approach the feasibility of redox reactions involving potentially organic and inorganic compounds can be evaluated in aquifers. That does, however, not imply that the corresponding reactions also occur within the relevant time scale. For this the kinetics of the reaction have to be known and have been studied for many reactions of potential relevance in aquifer systems. However, the kinetics of redox reactions are not subject of this section.

References

Sparks, D. (2002). Environmental Soil Chemistry, Second Edition, Academic Press, Chapters 5 and 8, ISBN 978-0126564464.

Essington, M.E. (2004). Soil and Water Chemistry: An Integrative Approach, Chapters 7 and 9, CRC Press, ISBN 978-0849312588

3.1.6. Question 1

Which processes control the movement of chemicals in groundwater? Define each of these processes.

3.1.6. Question 2

Redox reactions of contaminants in groundwater are controlled by the redox potential. What are the most important chemicals and processes determining the redox potential?

3.1.6. Question 3

How can the redox potential in aquifer be determined? What are the disadvantages of these methods?

3.1.7. Biota

(draft)

Author: Steven Droge,

Reviewer: Nico van der Brink, John Parsons

Leaning objectives:

You should be able to:

- explain the effect of cell and body composition on toxicokinetics of compounds
- describe the role of biota in the environmental fate of chemicals

Keywords: cellular composition, body composition, exposure routes, absorption, distribution

Introduction

Just like soil, water, and air, the organic tissue of living organisms can also be regarded as a compartment of the ecosystem where chemical pollutants can accumulate or can be broken down. The internal concentration in living organisms provide important information on chemical exposure and ultimately determines the environmental risk of pollution, but it is important to understand the key features of tissue that influence chemical partitioning into organisms. Chemical accumulation in the tissue of living organisms is a series of chemical and biological processes, briefly based on:

- chemical uptake (mostly permeation from bulk media over certain membranes into cells);

- internal distribution (e.g. via blood flows through organs);

- metabolism (e.g. biotransformation processes in for instance the liver).

- excretion (e.g. through urine and feces, but also via gills, sweat, milk, or hairs)

These four processes are the basis of toxicokinetic modeling, and are often summarized as Absorption, Distribution, Metabolism, and Excretion, or "ADME". These ADME processes can strongly vary for different polluting compounds due to the properties of the chemical structure. These ADME processes can also strongly vary for different organisms, because of:

- the physiological characteristics (e.g. having gills, lungs, or roots, availability of specific chemical uptake mechanisms, presence of specific metabolic enzymes, size-related properties like metabolic rate),

- the position in the polluted environment (flying birds or midge larvae living in sediment),

- the interaction with the polluted environment (living in soil or water, food choice, etc.)

- the behaviour in the polluted environment (being sessile or able to move (temporarily) away from a polluted spot).

More details of these toxicokinetic processes are presented in section 4.1 on Toxicokinetics and bioaccumulation. The current module aims to provide a summary of the key features of different tissue components that explain the internal distribution of chemicals (distribution), the different types of contact between pollutants and organisms (exposure-absorption), and temporal changes in physiology that may affect internal exposure (e.g. excretion, which includes examples such as release of POPs via lactation, and increasing POP concentrations during starvation). Before we discuss how chemicals are taken up into biota, it is important to first define the key chemical properties and the molecular composition of tissue that influence the way chemicals are absorbed from the surrounding environment and distributed throughout an organism.

Absorption-distribution: Tissue building blocks

All organisms are composed of cells, which are composed of a cell membrane, surrounding the largely watery solution filled with inner organelle membranes, protein structures, and DNA/RNA. Prokaryote organisms such as bacteria, but also algae, fungi and plants have reinforced membranes with cell walls to prevent water leaking by high osmotic pressures, and to protect the cell

membrane. Metabolic energy is stored in large molecules such as fatty esters and sugars. Remarkably, for all existing living organism species, these tissue components are mostly structures made out of relatively simple and repetitive molecular building blocks, with minor variations on side chains. See examples in Figure 1. The composition of organs, as a collection of specific cells, in terms of the percentage of lipids, proteins and carbohydrates is important for the overall toxicokinetics of chemicals in the whole organism.

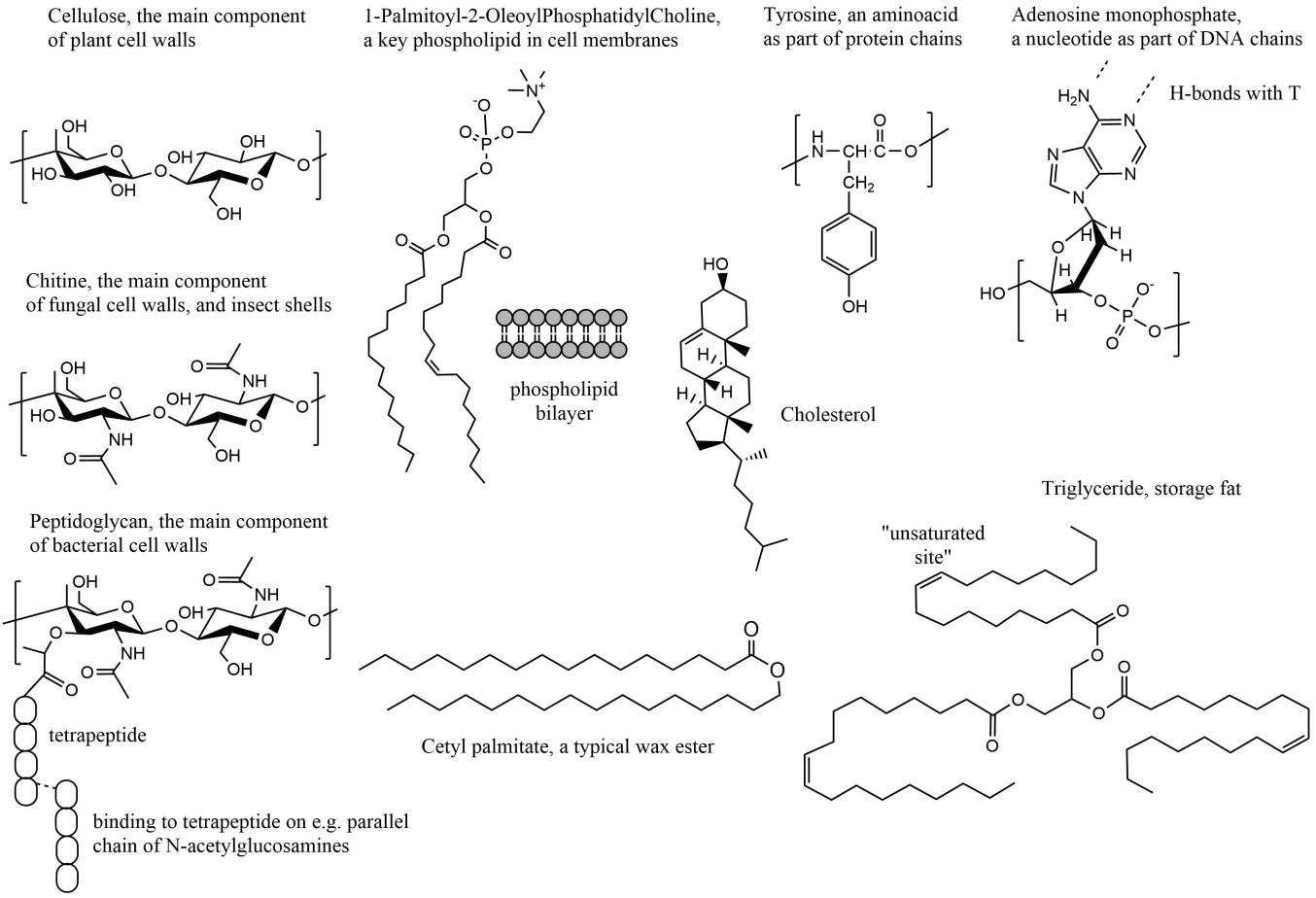


Figure 1. the key molecular components and partial building blocks of living tissue.

Cell walls are mostly made from highly polar polysaccharides, e.g.:

- cellulose, a polymer of sugary molecules, and chitin in fungi, which is highly polar and therefore permeable for water.
- peptidoglycan semi-crystal structures surrounding bacteria (90% of the dry weight of Gram-positive bacteria but only 10% of Gram-negative strains), a mixed polymer between N-acetylglucosamine (alike chitine) and short interconnecting 4 or 5 amino acid chains.
- lignin, a polar (~30% oxygen) but more hydrophobic supra-structure of polymerized phenolic molecules lining the main plant vessels that transport water.

The specific algae group of diatoms have a cell wall composed of biogenic silica (hydrated silicon dioxide), typically as two valves that overlap each other surrounding the unicellular species. Diatoms generate about 20 percent of the oxygen annually produced on the planet, and contribute nearly half of the organic material found in the oceans. With their specific cell wall structure, diatoms take in over 6.7 billion metric tons of silicon each year from the waters in which they live, which creates huge deposits when they die off.

Cell membranes are made up mostly of a phospholipid bilayer, with each phospholipid molecule basically having a polar and ionized headgroup connected to two long alkyl chains (Figure 1 example with POPC type phospholipid). The outer sides of a phospholipid bilayer are hydrophilic (water-loving), the inside is hydrophobic (water-fearing). Ions (inorganic salts, nutrients, metals, strong acids and ionized biomolecules) do not readily permeate through such a membrane passively, and require specific

transport proteins that can transport as well as regulate ions in and out of the cell interior. Cholesterol molecules stabilize the fluidity of the membrane bilayers in cells of most organisms, but for example not in most Gram negative bacteria. Dissolved neutral chemicals may passively diffuse through phospholipid bilayers into and out of cells.

Proteins are chains of a variety of amino acids, 21 of which are known to be genetically coded, and of which humans can only produce 12. The other nine must be consumed, and are therefore called essential amino acids (coded H, I, L, K, M, F, T, W, V). Proteins form complex 3 dimensional structures that allow for enzymatic reactions to occur effectively and repeatedly. There are two amino acids with side chains that carry a positive charge at neutral pH: Arginine (pKa 12) and Lysine (pKa 10.6), and two amino acids with side chains that carry a negative charge at neutral pH: Aspartic acid (pKa 3.7) and Glutamic acid (pKa 4.1). Some amino acids carry typical hydrophobic side chains: amongst others Leucine and Phenylalanine. Cysteine has a thiol (SH) moiety that can form strong connective disulfide interactions with spatially nearby cysteine side groups in the 3D structure. The key blood transport protein albumin, for example, contains about 98 anionic amino acids, and 83 cationic amino acids, and about 35 cysteine residues.

DNA and other genetically encoding chains are composed of 4 different nucleotides that form a double helix of two opposing strands, held together by hydrogen bonds connecting the complementary bases: A and T (or A and U in RNA), and G and C. DNA can be densely packed around histone proteins, and is either or part of the cellular cytoplasm (in prokaryotic species) or separated within a membrane (in eukaryotic species). DNA is not a critical accumulation phase for chemicals, but of course it is a cellular structure where pollutants can strongly impact all kinds of cellular processes when they react with DNA components or affect the structural organisation otherwise.

Storage fat provides for many animals and fruits of plants an important energy reserve, but also insulates warm-blooded animals in cold climates, lubricates joints to move smoothly, and protects organs from shocks (e.g. eyes and kidneys). Seeds and nuts may contain up to 65% (walnuts) of fatty components, which of course provides energy for initial growth, but from which also oil can be pressed. Storage fat in most animals is present in the form of triglycerides, and as such neutral and very hydrophobic phases within tissue. Polyunsaturated fatty acid esters like omega-6 and omega-3 fatty acids are abundant in fish (eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA)) and in seeds and plants (mostly alpha-linolenic acid (ALA), but algae also contain EPA and DHA). The high intake of algae by fish in aqueous food webs based on algae, results in the high EPA/DHA levels in many fish species, as they mostly cannot make it themselves (https://www.pufachain.eu). Humans can make some EPA and DHA from ALA.

The average composition of living organisms based on the key tissue components lipid, protein, and carbohydrate can range widely, as illustrated in Table 1.

Table 1. Tissue structure composition of the average dry weight of different organisms.

Organism	lipid % of d.w.	protein % of d.w.	carbohydrate % of d.w.
Grass	**0.5-4**	**15-25**	**60-84**
Phytoplankton	**20**	**50**	**30**
Zooplankton	**15-35**	**60-70**	**10**
Oyster	**12**	**55**	**33**
Midge larvae	**10**	**70**	**20**
Army cutworms (moth larvae)	**72% of body**		
Pike filet	**3.7**	**96.3**	**0**
Lake trout	**14.4**	**85**	**0**
Eel (farmed for 1.5y)	**65**	**~34**	**1**
Deer game meat	**10**	**90**	**0**

Table 2. Estimates on the tissue structure composition of a woman (BW= 60 kg, H = 163 cm, BMI = 22.6 kg/m^2), (taken from Goss et al., 2018). Bones are not included.

Organ	Total organ volume (mL)	moisture content	phospholipid % of d.w.	storage lipid % of d.w.	protein % of d.w.
Adipose	22076	26.5%	0.3%	93.6%	6.1%
Brain	1311	80.8%	35.4%	22.6%	42.1%
Gut	1223	81.8%	9.9%	22.0%	68.1%
Heart	343	77.3%	19.1%	17.7%	63.2%
Kidneys	427	82.4%	16.5%	5.9%	77.7%
Liver	1843	79.4%	19.4%	8.0%	72.6%
Lung	1034	94.5%	13.1%	13.7%	73.2%
Muscle	19114	83.7%	2.6%	2.5%	95.0%
Skin	3516	71.1%	2.6%	22.9%	74.5%
Spleen	231	83.5%	5.6%	2.4%	92.0%
Gonads	12	83.3%	18.8%	0.0%	81.3%
Blood	4800	83.0%	2.5%	2.4%	95.1%
total	55929	60.2%	1.7%	71.0%	27.3%

Different organs in a single species can also largely differ in their composition, as well as their contribution to the overall body, as shown for a human in Table 2. Most organs have a moisture content >75%, but overall the moisture content is considerably lower, due to the low moisture content of bones and adipose tissue. Adipose tissue is by far the largest repository of lipids, but made up mostly of storage lipid, while the brain is also particularly rich in both lipids, but particularly enriched in phospholipids of cell membranes. Muscles and blood contain relatively high protein content.

Absorption-distribution: Chemical properties

The influence of chemical structure on the accumulation of chemicals in the biotic compartments is largely dependent on their bioavailability, as discussed in more detail in section 4.1 on Toxicokinetics and bioaccumulation, as well as on the basic binding properties as the results of the chemical's hydrophobicity and volatility (section 3.4 on Partitioning and partitioning constants) and ionization state (section 2.2.6 on Ionogenic organic chemicals). In brief, the more non-polar the composition of a chemical, the more hydrophobic it is, the higher its affinity to partition from dissolved phases (both externally as well as internally) into poorly hydrated tissue phases such as storage fat and cell membranes. For this reason, the main issue with classical organic pollutants such as dioxins, DDT, and PCBs, is often their high hydrophobicity which results in strong accumulation in tissue. Such chemicals often take very long times to excrete from the tissue if they are not made less hydrophobic via biotransformation processes. This leads to foodweb accumulation and specific acute or chronic toxic effects at a certain organism level (section 4.1.6 on Food chain transfer). Proteins and sugary carbohydrates are mostly comprised of extended series of polar units and thus strongly hydrated, and bind hydrophobic chemicals to a much lower extent. Proteins may have three dimensional pockets that could fit either hydrophilic of hydrophobic chemicals, and as such act as transport proteins in blood throughout the body (transporting fatty acids for example), based on the specific binding affinity. Many protein based receptors are also based on a specific binding affinity, and in many cases this involves (combinations of) polar and electrostatic interactions that also have an optimum three-dimensional fitting space. Volatile chemicals are more abundantly present in the gas phase rather than being dissolved, and are more readily in contact with biota via gas-exchange on the extensive surfaces of lungs of animals and leaves of plants.

In order to be taken up into cells, or into organs, chemicals have to permeate through membranes. For most organic pollutants, the passive diffusion through phospholipid bilayers has an optimum at a certain hydrophobicity. The high accumulation in the membrane ensures desorption into the adjacent cellular solution. It is assumed that ionized chemicals have a passive permeation rates that are either negligible or at least orders of magnitude lower than that of corresponding neutral chemicals. For this reason, all kinds of molecular intra-extracellular gradients can be readily maintained, for example for protons (H^+) or sodium-potassium (Na^+/K^+). The movement of very polar and ionic chemicals can be tightly regulated by transport proteins protruding the membrane bilayer. Specific molecules can be actively excreted from cells (e.g. certain drugs) or reabsorbed (e.g. in the kidneys back into the blood stream). This again is based on three dimensional fitting in the transport pocket and stepwise movement through the protein structure, and costs energy. For acids and bases with a very small fraction of neutral species at physiological pH, the passive permeation over membranes may still be dominated by the neutral species.

Exposure: Contact between biota and various environmental compartments

There are multiple routes by which chemicals can enter the tissue of biota, for example via respiratory organs, through digestion of contaminated food, or dermal contact. Most animals need to take in a more or less constant flux of oxygen and water, and periodically food, to release nutrients and energy from the food. Of course they also need to release CO_2 (and other chemicals) as waste. Pollutant chemicals are taken up alongside these basic processes, and it depends on the chemical properties and the efficiency of the uptake route how much the organism will take in from these different exposure routes.

Plants need plenty of water and during daytime photosynthesis need CO_2, but also require oxygen during the night. High algal densities can deplete the oxygen levels in shallow aquatic systems during the night, and replenish oxygen levels during daytime. Oxygen is plentiful in air (200,000 parts per million in the air), but it is considerably less accessible in water (15 parts per million in cool, flowing water), and often depleted below the first few mm of sediment. To obtain sufficient oxygen, water and food, aquatic organisms have to pass large volumes of waters through their gills. Sediment-dwelling organisms either have hemoglobin to bind oxygen, or constantly pump fresh overlying water through burrows created in sediment, often lined with mucus. Living organisms are thus constantly in contact with water dissolved pollutants, and air breathing organism are readily exposed to air pollutants. To simplify the domain of living organisms as part of this module about the biotic compartment and how they get into contact with chemicals relevant to Environmental Toxicology, they can for example be divided in:

- , which often take in large amounts of water from their surroundings via their roots, driven by the evaporation of water at the leaves and resulting internal flows.
- **water breathing organisms**, which pass large amounts of water through gills or gill like structures (tubules or other thin skin structures close to where water is passing) in order to take in enough oxygen and reduce the built up of CO_2. Filter feeders like oysters and mussels, that populate enormous surfaces as reefs or banks, can turnover a huge volume of water on a daily basis, and thus allow dissolved chemicals get in close contact to the outer membranes.
- **air breathing organisms**, which can effectively exchange large quantities of volatile and gaseous chemicals with the air (oxygen but also organic compounds), but typically take in less/non-volatile chemicals via food and require active excretion and metabolism for the emission of less/non-volatile chemicals.

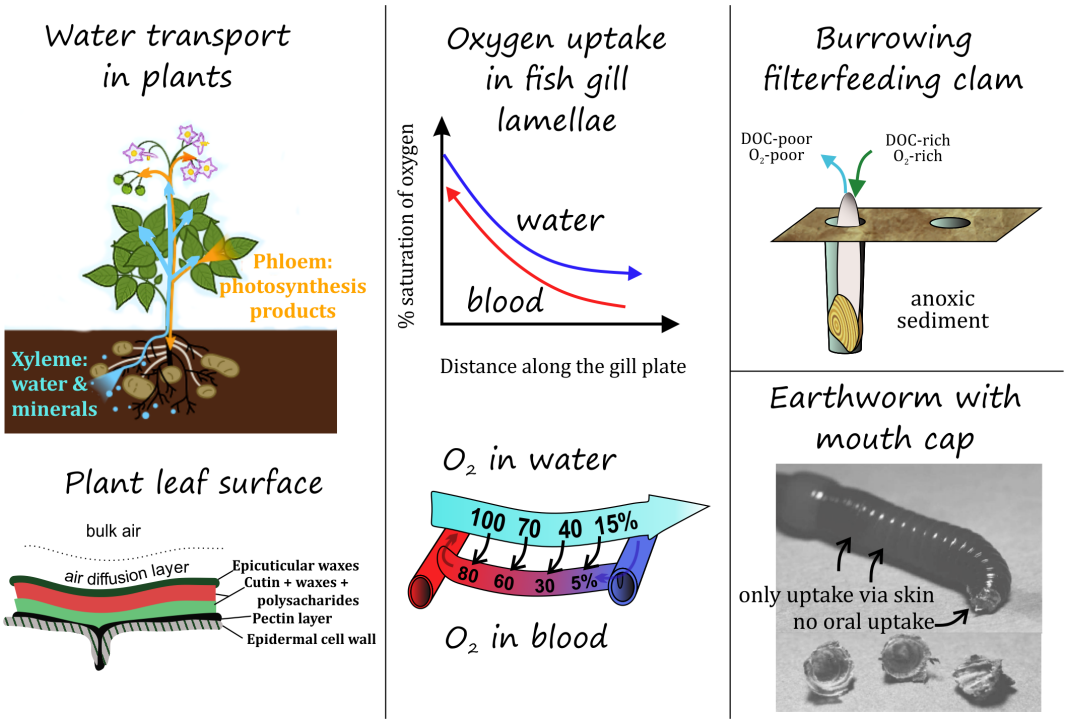


Figure 2. Left panel: Plant water transport (adapted from https://en.Wikipedia.org/wiki/Xylem), a schematic view of the waxy upper layers on leaves (adapted from Moeckel et al. 2008)). Middle panel: water flushing along gill lamellae replenishes oxygen in blood. **Top Right panel:** Many benthic organisms living in anoxic sediments require overlying oxygen rich water, and feed on dissolved particles. **Lower Right Panel:** Earthworms take up oxygen via diffusion through their skin, but also metals and organic pollutants, as shown in studies where the mouth part has been glued shut during exposure (adapted from Vijver et al., 2003).

Plants

Nearly all plants have roots below ground, a sturdy structure of stem and branches above ground, and leaves. Along with soil pore water, soluble chemicals are readily transported from roots of the plant in the internal circulation stream through xylem cells, which are lined with water impenetrable lignin (see Figure 2). Moderately hydrophobic chemicals (K_{ow} of 1-1000) are rapidly transported from roots to shoots to leaves, while hydrophobic chemicals may be strongly retained on the membranes and cell walls and mostly accumulate in root sections, limiting transport to above-ground plant tissues. Roots may also actively release considerable quantities of chemicals to influence the immediate surrounding media of the roots (rhizosphere), e.g. to stimulate microbial processes or pH in order to release nutrients. These plant root 'exudates' can be ions, small acids, amino acids, sterols, etc. Chemicals that enter plants via leaves, such as pesticides or semi-volatile organic pollutants, can be redistributed to other plant parts via the phloem streams.

The transport through xylem up to higher plant tissues occurs via capillary forces and is enhanced by three passive phenomena:

- the high sugar content in the phloem causing osmotic pressure to attract water from other parts.
- the evaporation of water creates surface tension on thousands of cells that pulls water from the soil through the xylem system.
- the osmotic pressure of root cells compared to soil pore water. Root pressure is highest in the morning before the stomata open and allow transpiration to begin.

As a result of the capillary forces needed to pull water up against gravity, and a certain maximum diameter of the vessels to do so, there is a maximum possible plant height of 122-130 m (Koch et al., 2004) which compares to Redwood trees (Sequoias) reaching a maximum height of 113 m.

Most leaves are covered with a waxy layer, to prevent damage and water evaporation. This wax layer may be 0.3-4.6 µm thick (Moeckel et al., 2008). Large forests provide enormous hydrophobic surfaces to which semi-volatile organic chemicals (SVOC) can bind out of air, which influences the global distribution of chemicals such as PCBs. Partitioning of SVOCs on the vegetation of extended grasslands contaminates the base of the food chain, as well as agricultural and cattle sectors used by humans. The grass/corn-cattle-milk/beef food chain accounts for the largest portion of background exposure of the European and North

American population to many persistent SVOCs. The absorption rate of chemicals on the leaves often also depends on the air boundary layer surrounding leaves, which limits diffusion into the leave surfaces. Of course, all kinds of other factors such as wind speed, canopy formation and cuticle thickness also control exchange between leaves and gas phase (see also section 3.1.2 on the Atmosphere). Tiny openings or pores on the lower side of the leaves, called stomata, allow furthermore for gas exchange. In warm conditions, stomata can close to prevent water evaporation, but gas exchange is needed in many plant types to allow for CO_2 to be metabolized and the release of O_2 that is produced. The waxy layer on leaves can trap gaseous organic chemicals. Many plants, like coniferous trees, produce resins to provide effective defense against insects and diseases, and these resins release large amounts and structurally highly diverse organic volatiles such as terpenes and isoprenes (Michelozzi, 1999). These plant-produced volatiles can even contribute to ozone formation. Plants thus accumulate chemicals from their environment, but also release chemicals into the environment. It thus also matters for the exposure of grazing organisms to certain types of pollutants whether they eat roots, shoots, leaves, seeds or fruits of plants living in contaminated environments.

Organisms using dissolved oxygen

The 'gill' movements of water breathers create a constant flux of chemicals dissolved in bulk water along outer cell membranes (or mucus layers surrounding cell membranes) of gills. A 1 kg rainbow trout fish ventilates about 160 mL/min, so 230 L/day (Consoer et al., 2014). The total gill surface area in a fish depends on species behaviour and weight (active large fish require a lot of oxygen), and equals to about 1-6 cm^2 per g fish (Palzenberger & Pohla, 1992). For a 1 kg fish of ~20 cm length, the ~1000 cm^2 gill area compares to a ~500 cm^2 outer body surface. This results in an effective partitioning of chemicals between water and cell membranes. Within the gills, the cells are in close contact with the blood system of the organism, and the build-up of chemical concentrations in the outer cells provides an effective exchange with the blood stream (or other internal fluids, Figure 2) flushing along that redistribute chemicals to the inner organs. The reverse equally occurs: chemicals dissolved in blood stream coming from organs will also rapidly exchange with bulk external water if concentrations are lower.

Of course, many pollutants can also enter water breathing organisms via food, but the gills-water exchange is very effective in controlling the distribution of chemicals. The salinity of water plays a strong role in the need of water breathing organisms to "drink" water, and hence take in contaminants via this route.

BOX 1. Osmoregulation (MSc level)
Most aquatic vertebrate animals are osmoregulators: their cells contain a concentration of solutes that is different than the water around them. Fish living in freshwater typically have a cellular osmotic level (300 millOsmoles per Liter, mOsm/L) that is higher than the bulk fresh water (~20-40 mOsm/L), so a lot of water flows passively via the gills (not via the skin) into the tissue of fish. They are thus constantly taking in water (water molecules only) via the gills, which needs to be controlled, e.g. by strongly diluting the urine. They do also take in some water in their gastro-intestinal tract (GIT). Marine fish have similar cellular osmotic levels as freshwater fish, but the salty water (1000 mOsm/L) causes water to move out of the gill tissue through the linings of the fish's gills by osmosis, which needs to be replenished by active intake of salty water, and separate excretion of the salts. Most invertebrate organisms in oceans have an internal overall concentration of dissolved compounds comparable to the water they live in, so that they don't suffer from strong osmotic pressures on their soft tissue (osmoconformers).

A single adult oyster can cleanse about 200 liters of water per day (https://www.cbf.org/about-the-bay/mo...act-sheet.html). Plans to re-populate the harbour of New York with 1 billion oysters on artificial substrates can have enormous impacts on chemical redistribution. A single 2 cm zebra mussel (*Dreissena polymorpha*) that inhabits the shallow Lake IJsselmeer ($6.05x10^{12}$ L) in can filter about 1 L per day, and the high densities of these and related species in this fresh water lake can turn over the lake volume once or twice per month (Reeders et al., 1989).

Even many soil organism are constantly in contact with wet soil surfaces, and contact between soil pore water and the outer surfaces (gills/soft skin areas) dominates the routes of chemical exchange for many chemicals. Earthworms, for example, do not have lungs and they exchange oxygen through their skin. Earthworms eat bacteria and fungi that grow on dead and decomposing organic matter, and thus act as major organic matter decomposers and recycling of nutrients. Earthworms dramatically alter soil structure, water movement, nutrient dynamics, and plant growth. It is estimated that earthworms turn over the top 15 cm of soil in ten to twenty years (LINK), and so they also are able to mix surface bound pollution into a substantial soil layer. In terms of biomass, earthworms dominate the world of soil invertebrates, including arthropods. In order to better understand how much contamination earthworms take in via food or via their skin, several studies have used earthworms in exposure tests with part of the

organisms having their mouth parts sealed with surgical glue (Vijver et al., 2003). Uptake rates of the metals Cd, Cu and Pb in sealed and unsealed earthworms exposed to two contaminated field soils were similar (Vijver et al., 2003), indicating main uptake through the skin of the worms. The uptake rates as well as the maximum accumulation level for several organic contaminants from artificially contaminated soil were also comparable between sealed and non-sealed worms (Jager et al., 2003). The dermal route is thus a highly important uptake route for organic chemicals too. Dermal uptake by soil organisms is generally from the pool of chemicals in the soil pore water, hence the distribution of chemicals between solid particles and organic materials in the soil and the soil pore water is extremely important in driving the dermal uptake of chemicals by earthworms (See section 3.4 on Partitioning and partitioning constants, section 3.5 on Metal speciation and 3.6 on Availability and bioavailability).

Air breathing organisms

Air breathing organisms typically take in less/non-volatile chemicals via food and require active excretion and metabolism for the elimination of these chemicals via e.g. feces and urine. Dermal uptake is generally assumed to be negligible, while the intake of chemicals via air particles can be relatively high, for example of pollutants present in house dust, or contaminants on aerosols. The excretion via air particles is clearly not a dominant route. The chemicals in the food matrix inside the gastro-intestinal tract often first need to be fully dissolved in gut fluids, before they can pass the mucus layers and membranes lining the gastrointestinal tract and enter blood streams for redistribution. However, 'endocytosis' may also result in the uptake of (small) particulate chemicals in cells by first completely surrounding the particle by the membrane, after which the encapsulating membrane buds buds off inside the cell and forms a vesicle. Notwithstanding this endocytosis, chemical fractions of pollutants strongly sorbed to non-digestible parts may not always automatically be taken up from food. Grazing animals typically require microbial conversion in their gut to digest plant material like cellulose and lignin into chemical components that can be taken up as energy source.

Many aquatic foodwebs are structured such that they begin with aquatic plants being eaten by water breathing organisms, with air breathing marine animals or birds on top of the food chain. These air breathing top predators take in pollutants largely through their diet, but lack the effective blood-membrane-water exchange through gills. The blood-membrane-air partitioning in lungs is far less effective in removing chemicals via passive partitioning. For this reason, many top predators of foodwebs have the highest concentrations of pollutants. The chemical distribution in foodwebs will be discussed in more detail in section 4.1.6.

References

Palzenberger & Pohla 1992, Reviews in Fish Biology and Fisheries, 2, 187-216.

Jager, T.; Fleuren, R.H.L.J.; Hogendoorn, E.A.; de Korte, G. 2003. Elucidating the routes of exposure for organic chemicals in the earthworm, *Eisenia andrei* (Oligochaeta). Environ. Sci. Technol. 37 (15), 3399-3404

Koch et al. 2004, Nature (428) 851-854.

Moeckel et al. 2008, Environ. Sci. Technol. 42, 100-105.

Michelozzi 1999, Defensive roles of terpenoid mixtures in conifers, Acta Botanica Gallica 146 (1), 73-84.

Consoer et al. 2014, Aquatic Toxicology 156, 65-73.

Reeders et al. 1989, Freshwater Biology 22 (1), 133-141.

Vijver et al. 2003, Soil Biology and Biochemistry 35, 125-132.

Goss et al. 2018, Chemosphere 199, 174-181.

3.1.7. Question 1

Explain whether a lean fish of 1 kg likely contains a lower concentration of hydrophobic pollutant as a 1 kg fatty fish caught in the same lake.

3.1.7. Question 2

Explain how bivalves (clams, muscles, oysters) can reduce contamination with a hydrophobic chemical from a lake

---

3.1.7. Question 3

In many countries, the levels of water pollution have been strongly reduced compared to 30 years ago, but a lot of sediments are still contaminated with historic pollutants that strongly bind to the sediment particles. Explain how historic contamination in a sediment can be brought into suspension by benthic organisms.

3.1.7. Question 4

Discuss whether the internal concentrations of very hydrophobic chemicals in fatty fish gets higher or lower during an extended period of starvation, e.g. during migration towards breeding grounds. Note that very hydrophobic chemicals may have elimination half lives of months or even longer.

3.1.7. Question 5

Discuss whether the internal concentrations of very hydrophobic chemicals in female organisms gets higher or lower during a reproduction cycle, e.g. by generating offspring and or lactation. Note that very hydrophobic chemicals may have elimination half lives of months or even longer.

This page titled 2.3.3.1: Environmental compartments is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by John Parsons & Steven Droge via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1: Environmental compartments** by John Parsons, Steven Droge is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.3.2: Sources of chemicals

3.2. Sources of chemicals

Under review

This page titled 2.3.3.2: Sources of chemicals is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **3.2: Sources of chemicals** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.3.3: Pathways and processes determining chemical fate

3.3. Pathways and processes determining chemical fate

Authors: Dik van de Meent, Michael Matthies

Reviewer: John Parsons

Leaning objectives:

You should be able to

- explain how emission of chemicals into the environment leads to exposure of ecosystems, populations, and organisms including man.
- understand how quantitative knowledge of process kinetics is of use in mathematical modeling of environmental fate
- understand the concept of characterization of fate processes in terms of rate constants and half-lives
- give examples of relevant fate processes and briefly describe them

Keywords: fate processes, degradation, transport, partitioning

Characterizing 'fate'

Chemicals can escape during all steps of their life cycle, e.g. manufacturing, processing, use, or disposal. Release of chemicals into the environment necessarily leads to exposure of ecosystems, populations, and organisms including man. Exposure assessment science seeks to analyze, characterize, understand and (quantitatively) describe the pathways and processes that link releases to exposure. Chemicals in the environment undergo various transport, transfer and degradation processes, which can be described and quantified in terms of loss rates, i.e. the rates at which chemicals are lost from the environmental compartment into which they are emitted or transferred from adjacent compartments. Exposure assessment science aims to capture the 'environmental fate' of chemicals in process descriptions that can be used in mass balance modeling, using mathematical expressions borrowed from thermodynamic laws and chemical reaction kinetics (Trapp and Matthies, 1998).

The 'fate' of a chemical in the environment can be viewed of as the net result of a suite of transport, transfer and degradation processes (see Section 3.4 on partitioning and partitioning constants, Section 3.6 on availability and bioavailability, Section 3.7 on degradation) that start to act on the chemical directly after its emission (see Section 3.2 on sources of emission) and during the subsequent environmental distribution. Environmental fate modeling (see Section 3.8 on multimedia mass balance modelling) builds on this knowledge by implementing the various degradation, transfer and transport processes derived in exposure assessment science in mathematical models that simulate 'fate of chemicals in the environment'.

First-order kinetics

In chemical reaction kinetics, the amount of chemical in a 'system' (for instance, a volume of surface water) is described by mass balance equations of the kind:

$$\frac{dm}{dt} = i - k\,m$$ (eq. 1)

where $\frac{dm}{dt}$ is the rate of change (kg.s^{-1}) of the mass (kg) of chemical in the system over time t (s), i is the input rate (kg.s^{-1}) and k (s^{-1}) is the reaction rate constant. Mathematically, this equation is a **first-order differential equation** in m, meaning that the loss rate of mass from the systemis proportional to the **first power** of m. Equation 1 is widely applied in description and characterization of environmental fate processes: environmental fate processes generally obey first-order kinetics, and can generally be characterized by a first-order reaction rate constant k^{1st}:

$$\frac{dm}{dt} = -k^{1st} m$$ (eq. 2)

Such loss rated equations can also be formulated in integral format, which is obtained by integration of equation (2) over time t with initial mass m_0 = m(0):

$$m_t = m_0\, e^{-k^{1st}\, t}$$ (eq. 3)

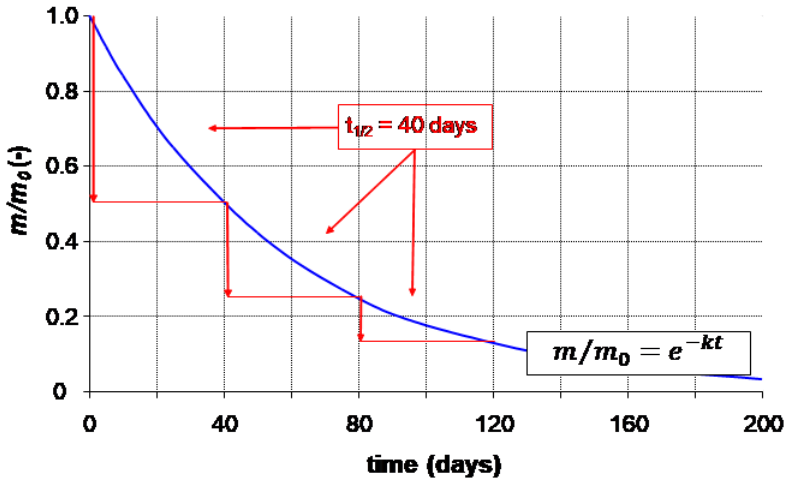


Figure 1. Graphical representation of equation 3.

Decrease of relative mass of a chemical in an environmental compartment would follow the blue curve when the loss process is given by a first-order differential loss equation. Loss processes that obey first-order kinetics have constant half-lives (here time $t_{1/2}$ *=40 days).*

As shown in Figure 1, first-order loss processes are expected to result in exponential decrease of mass from which concentration can be calculated by dividing m with the compartment volume. Using the value $t_{1/2}$ for *t* in equation 3, it follows directly that the value of $t_{1/2}$ is inversely proportional to the first-order loss rate constant k^{1st}:

$$\frac{m_t}{m_0} = \frac{1}{2} = e^{-k^{1st} t_{1/2}} \rightarrow ln(\frac{1}{2}) = -k^{1st} t_{1/2} \rightarrow t_{1/2} = \frac{ln(2)}{k^{1st}} \quad \text{(eq. 4)}$$

which shows that half-life time constant, i.e. independent of the concentration of the chemical considered. This is the case for all environmental loss processes that obey first-order kinetics. First-order loss processes can therefore be sufficiently characterized by the time required for disappearance of 50% of the amount originally present.

The disappearance time *DT50* is often used in environmental regulation but is only identically with the half-life if the loss process is of first order. Note that the silent assumption of constancy of half-life implies that the process considered is assumed to obey first-order kinetics.

Abiotic chemical reactions

Occurrence of true first-order reaction kinetics in chemistry is rare (see Section 3.7 on degradation). It occurs only when substances degrade spontaneously, without interaction with other chemicals. A good example is radio-active decay of elements, with a reaction rate proportional to the (first power of) the concentration (mass) of the decaying element, as in equation 3.

Most chemical reactions between two substances are of second order:

$$\frac{dm}{dt} = -k^{2nd} m_1 m_2 \quad \text{(eq. 5)}$$

or, when a chemical reacts with itself:

$$\frac{dm}{dt} = -k^{2nd} m_1^2 \quad \text{(eq. 6)}$$

because the reaction rate is proportional to the concentrations (masses) of both of the two reactants. It follows directly from equation 2. As the concentrations (masses) of both reactants decrease as a result of the reaction taking place, the reaction rate decreases during the reaction, more rapidly so at high initial concentrations. When second-order kinetics applies, half-life is not constant, but increases with ongoing reaction, when concentrations decrease. In principle, this is the case for most chemical reactions, in which the chemical considered is transformed into something else by reaction with a transforming chemical agent.

In the environment, the availability of second reactant (transforming agent) is usually in excess, so that its concentration remains nearly unaffected by the ongoing transformation reaction. This is the case, for oxidation (reaction with oxygen) and hydrolysis (reaction with water). In these cases, the rate of reaction decreases with the decreasing concentration of the first chemical only:

$$\frac{dm}{dt} = -(k^{2nd}\, m_2)\, m_1 = k^{1st}\, m_1 \quad \text{(eq. 7)}$$

and reaction kinetics become practically first-order: so-called *pseudo* first-order kinetics. Pseudo first-order kinetics of chemical transformation processes is very common in the environment.

Biotic chemical reaction

Chemical reactions in the biosphere are often catalyzed by enzymes. This type of reaction is saturable and the kinetics can be described by the Michael-Menten kinetic model for single substrate reactions. At low concentrations, there is no effect of saturation of the enzyme and the reaction can be assumed to follow (pseudo) first order kinetics. At concentrations high enough to saturate the enzyme, the rate of reaction is independent of the concentrations (masses) of the reactants, thus constant in time during the reaction, and the reaction obeys zero-order kinetics. This is true for catalysis, where the reaction rate depends only on the availability of catalyst (usually the reactive surface area):

$$\frac{dm}{dt} = -k^{zero}\, m^0 = \text{constant.} \quad \text{(eq. 8)}$$

One could say that the rate is proportional to the zero-th power of the mass of reactant present. In case of zero-order kinetics, the half-life times are longer for greater initial concentrations of chemical.

An example of zero-order reaction kinetics is the transformation of alcohol (ethanol) in the liver. It has been worked out theoretically and experimentally that human livers remove alcohol from the blood at a constant rate, regardless the amount of alcohol consumed.

Microbial degradation

Microbial degradation (often referred to as biodegradation) is a special case of biotic transformation kinetics. Although this is an enzymatically catalysed process, the microbial transformation process can be viewed of as the result of the encounter of molecules of chemical with microbial cells, which should result in apparent second-order kinetics (first order with respect to the number of microbial cells present, and first order with respect to the mass of chemical present):

$$\frac{dm}{dt} = -(k^{2nd}\, m_{bio})\, m = -k_{deg}\, m \quad \text{(eq. 9)}$$

where m_{bio} stands for the concentration (mass) of active bacteria present in natural surface water, and k_{deg} represents a pseudo-first-order degradation rate constant.

Advective and dispersive transport

Chemicals can be moved from one local point to another by wind and water currents. Advection means transport along the current axis whereas dispersion is the process of turbulent mixing in all directions. Advective processes are driven by external forces such as wind and water velocity, or gravity such as rain fall and leaching in soil. In most exposure models these processes are described in a simplified manner, e.g. the dispersive air plume model. An example of a first-order advective loss process is the outflow of a chemical from a lake:

$$\frac{dm}{dt} = -\frac{Q}{V} m = -k_{adv}\, m \quad \text{(eq. 10)}$$

where Q stands for the flow rate of lake water [m^3/s] and V is the lake volume [m^3]. Q/V is known as the renewal rate constant k_{adv} of the transport medium, here water. More sophisticated hydrological, atmospheric, or soil leaching models consider detailed spatial and temporal resolution, which require much more data and higher mathematical computing effort (see sections 3.1.2 and 3.8.2).

Transfer and partitioning

Due to Fick's first law the rate of transfer through an interface between two media (e.g. water and air, or water and sediment) is proportional to the concentration difference of the chemical in the two media (see section 3.4 on partitioning, and Schwarzenbach et al., 2017 for further reading). As long as the concentration in one media is higher than in the other, the more molecules are likely to pass through the interface. Examples are volatilization of chemicals from water (to air) and gas absorption from air (to water or soil), adsorption from water (to sediments, suspended solids and biota) and desorption from sediments and other solid surfaces.

When two environmental media are in direct contact, (first-order) transfer can take place in two directions, in the case of water and air by volatilization and gas absorption: Each at a rate proportional to the concentration of chemical in the medium of origin and each with a (first-order) rate constant characteristic of the physical properties of the chemical and of the nature of the interface (area, roughness). This is known as physical intermedia partitioning (see section 3.4 on partitioning), usually represented by a chemical reaction formula:

$$[M]_{water} \underset{\leftrightarrow}{\overset{exchange}{}} [M]_{air} \quad \text{(eq. 11)}$$

where [*M*] stands for a (mass) concentration (unit mass per unit volume) and the double arrow represents forward and reverse transport. Intermedia partitioning proceeds spontaneously until the two media have come to thermodynamic equilibrium. In the state of equilibrium, forward and backward rates (here: volatilization from water to air and gas absorption from air to water) have become equal. At equilibrium, the total (Gibbs free) energy of the system has reached a minimum: the system has come to rest, so that

$$k_{abs}[M]_{air} = k_{volat}\,[M]_{water} \rightarrow K_{AW} = \frac{[M]_{air}}{[M]_{water}} = \frac{k_{volat}}{k_{abs}} \quad \text{(eq. 12)}$$

and the ratio of concentrations of the chemical in the two media has reached its (thermodynamic) equilibrium value, called equilibrium constant or partition coefficient (see section 3.4 on partitioning).

Challenge

Challenge to environmental chemists is to describe and characterize the various processes of chemical and microbial degradation and transformation, of intra-media transport and intermedia transfer rate constants and of equilibrium constants, in terms of (i) physical and chemical properties of the chemicals considered and (ii) of the properties of the environmental media.

References

Schwarzenbach, R.P., Gschwend, P.M., Imboden, D.M. (2017). Environmental Organic Chemistry, Third Edition, Wiley, ISBN 978-1-118-76723-8.

Trapp, S., Matthies, M. (1998). Chemodynamics and Environmental Modeling. An Introduction. Springer, Heidelberg, ISBN 3-540-63096-1.

3.3. Question 1

Name and explain in your own words the essential environmentally relevant property of (pseudo) first-order kinetics.

3.3. Question 2

Give examples of transformation or transport processes that obey zero-order, first-order, second-order and pseudo first-order kinetics.

3.3. Question 3

Why is it useful to formulate environmental fate processes in terms of process rates and rate constants and equilibrium constants?

This page titled 2.3.3.3: Pathways and processes determining chemical fate is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **3.3: Pathways and processes determining chemical fate** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.3.4: Partitioning and Partitioning Constants

3.4. Partitioning and partitioning constants

3.4.1. Relevant chemical properties

Authors: Joop Hermens, Kees van Gestel

Reviewer: Steven Droge, Monika Nendza

Learning Objectives

You should be able to:

- define the concept of hydrophobicity and to explain which chemical properties affect hydrophobicity.
- define which properties of a chemical affect its tendency to evaporate from water.
- calculate fractions ionized for acids and bases.

Keywords: Hydrophobicity, octanol-water partition coefficients, volatility, Henry's Law constant, ionized chemicals

Introduction

Different processes affect the fate of a chemical in the environment. In addition to the transfer and exchange between compartments (air-water-sediment/soil-biota), also degradation determines the concentration in each of these compartments (Figure 1).

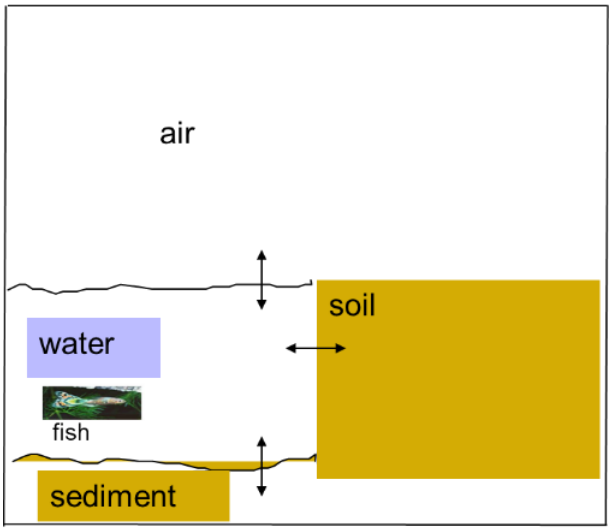


Figure 1. Environmental fate: exchange between compartments and degradation affect the concentration in each compartment.

Some of these processes are discussed in other sections (see sections on Sorption and Environmental degradation of chemicals). Some chemicals will easily evaporate from water to air, while others remain mainly in the aqueous phase or sorb to sediment and accumulate into biota.

These differences are related to only a few basic properties:

- Hydrophobicity (tendency of a substance to escape the aqueous phase)
- Volatility (tendency of a substance to vaporize)
- Degree of ionization

Hydrophobicity

Hydrophobicity means fear (phobic) of water (hydro). A hydrophobic chemical prefers to "escape from the aqueous phase" or in other words "it does not like to dissolve in water". Water molecules are tightly bound to each other via hydrogen bonds. For a chemical to dissolve in water, a cavity should be formed in the aqueous phase (Figure 2) and this will cost energy.

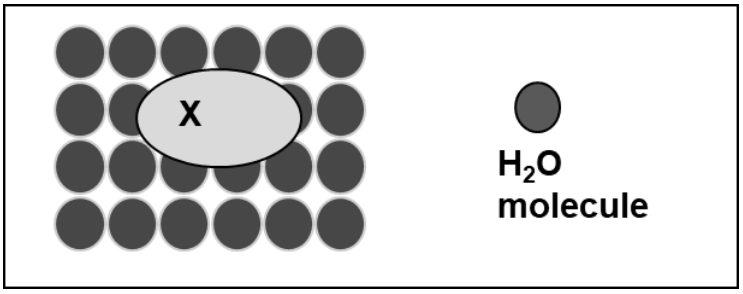


Figure 2. The formation of a cavity in water for chemical X.

Hydrophobicity mainly depends on two molecular properties:

- Molecular size
- Polarity / ability to interact with water molecules, for example via hydrogen bonding

It will take more energy for a chemical with a larger size to create the cavity making the chemical more hydrophobic, while interactions of the chemical with water will favour its dissolution making it less hydrophobic. Figure 3 shows chemicals with increasing hydrophobicity with increasing size and a decreasing hydrophobicity by the presence of polar groups (amino or hydroxy).

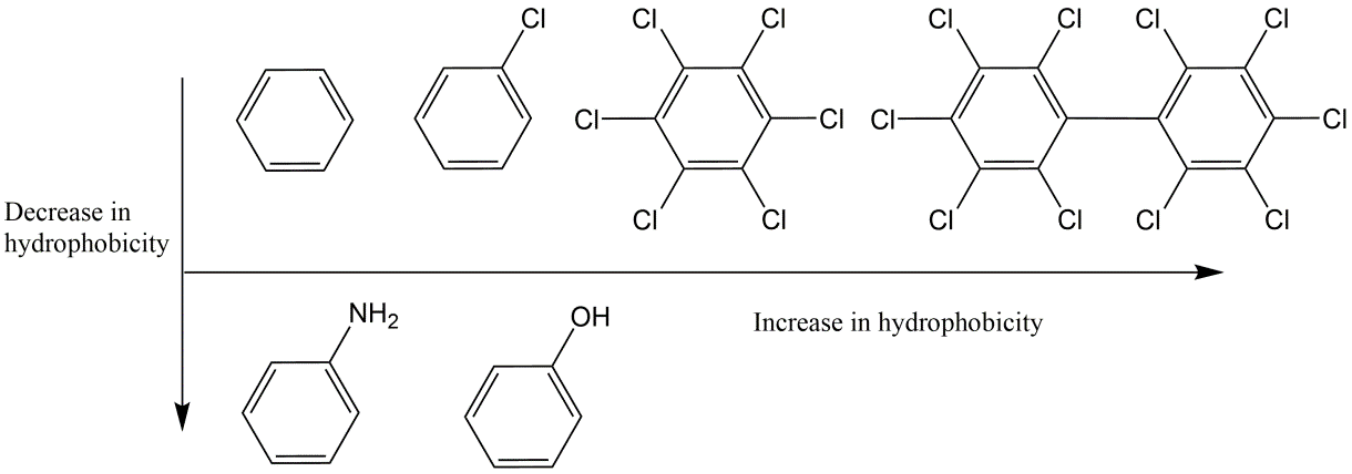


Figure 3. The effect of size and presence of polar groups on the hydrophobicity of chemicals. Increasing molecular size increases hydrophobicity; the introduction of polar groups leads to a decrease in hydrophobicity.

Most hydrophobic chemicals are non-polar organic micro pollutants. Well-known examples are the chlorinated hydrocarbons, such as polychlorinated biphenyls (PCBs) and polycyclic aromatic hydrocarbons (PAHs). Water solubility of these chemicals in general is rather low (in the order of a few ng/L up to a few mg/L).

The hydrophobic nature mainly determines the distribution of these chemicals in water and sediment or soil and their uptake across cell membranes. Additional Cl- or Br-atoms in a chemical, as well as additional $(CH)_x$ units, increase the molecular size, and thus a chemical's hydrophobicity. The increased molecular volume requires a larger cavity to dissolve the chemical in water, while they only interact with water molecules via VanderWaals interactions.

Polar groups, such as the -OH and -NH units on the aromatic chemicals in Figure 3, can form hydrogen-bonds with water, and therefore substantially reduce the hydrophobicity of organic chemicals. The hydrogen bonding of hydroxy-substituents works in two ways: The oxygen of -OH bridges to the H-atoms of water molecules, while the hydrogen of -OH can form bridges to the O-atoms of water molecules. Nearly all molecular units consisting of some kind of (carbon-oxygen) combination reduce the hydrophobicity of organic contaminants, because even though they increase the molecular volume they interact via hydrogen bonds (H-bonds) with surrounding water molecules. Additional polar groups in a chemical typically decrease a chemical's hydrophobicity.

Octanol-water partition coefficient:

A simple measure of the hydrophobicity of chemicals, originating from pharmacology, is the octanol-water partition coefficient, abbreviated as K_{ow} (and sometimes also called P_{ow} or P_{oct}): this is the ratio of concentrations of a chemical in n-octanol and in water, after establishment of an equilibrium between the two phases (Figure 4). The -OH group in n-octanol does allow for some hydrogen bonding between octanol-molecules in solution, and between octanol and dissolved molecules. However, the relatively long alkyl chain only interacts through VanderWaals interactions, and therefore the interaction strength between octanol-molecules is much smaller than that between water-molecules, and it is energetically much less costly to create a cavity to dissolve any molecule.

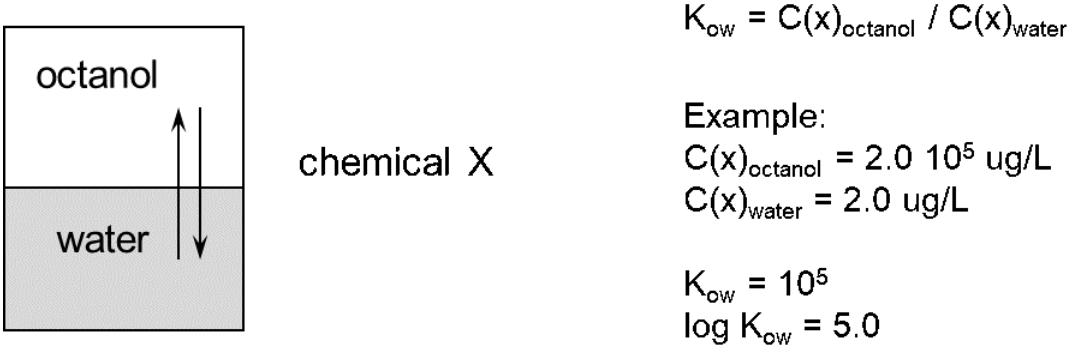


Figure 4. Distribution of chemical X between octanol and water and an example of a chemical with log K_{ow} of 5.0.

Experimentally determined K_{ow} values were used in pharmacological research to predict the uptake and biological activity of pharmaceuticals. Octanol was selected because it appears to closely mimic the nonionic molecular properties of most tissue components, particularly phospholipids in membranes. Since the beginning of the 1970s, K_{ow} values have also been used in environmental toxicology to predict the hazard and environmental fate of organic micro pollutants. Octanol may partially also mimic the nonionic molecular properties of most organic matter phases that sorb neutral organic chemicals in the biotic and abiotic environment.

Not unexpectedly, water solubility is negatively correlated with octanol-water partition coefficients.

In practice, three methods can be used to determine or estimate the K_{ow}:

Equilibration methods

In the shake-flask method (Leo et al., 1971) and the 'slow-stirring' method (de Bruijn et al., 1989), the distribution of a chemical between octanol and water is determined experimentally. For highly lipophilic chemicals (log K_{ow} > 5-6), the extremely low water solubility, however, hampers a reliable analytical determination of concentrations in the water phase. For such chemicals, these experimental methods are not suitable. During the last two decades, the use of generator columns has allowed for quantification of higher K_{ow} values. Generator columns are columns packed with a sorbing material (e.g. Chromosorb®) onto which an appropriate hydrophobic solvent (e.g. octanol) is coated that contains the compound of interest. In this way, a large interface surface area between the lipophilic and water phases is created, which allows for a rapid establishment of equilibrium. When a large volume of (octanol-saturated) water (typically up to 10 litres) is passed slowly through the column, an equilibrium distribution of the compound is established between the octanol and the water. The water leaving the column is passed over a solid sorbent cartridge to concentrate the compound and allow for a quantification of the aqueous concentration. In this way, it is possible to more reliably determine log K_{ow} values up to 6-7.

Chromatography

K_{ow} values may also be derived from the retention time in a chromatographic system (Eadsforth, 1986). The use of reversed-phase High Performance Liquid Chromatography (HPLC), thin-layer chromatography or gas chromatography results in a capacity factor (relative retention time; retention of the compound relative to a non-retained chemical species), which may be used to predict the chemical distribution over octanol and water. HPLC systems have shown most successful, because they consist of stationary and

mobile phases that are liquid. As a consequence, the nature of the phases can be most closely arranged to resemble the octanol-water system. Of course, this requires calibration of the capacity factors by applying the chromatographic method to a number of chemicals with well-known K_{ow} values. Chromatographic methods may reliably be applied for estimations of log K_{ow} values in the range of 2-8. For more lipophilic chemicals, also these methods will fail to reliably predict K_{ow} values (Schwarzenbach et al., 2003).

Calculation

K_{ow} values may also be calculated or predicted from parameters describing the chemical structure of a chemical. Several software programs are commercially available for this purpose, such as KOWWIN program of the US-EPA. These programs make use of the so-called fragment method (Leo, 1993; Rekker and Kort, 1979). This method takes into account the contribution to K_{ow} of different chemical groups or atoms in a molecule, and in addition corrects for special features such as steric hindrance or other intramolecular interactions (equation 1):

$$\log K_{ow} = \Sigma f_n + \Sigma F_p \quad \text{(eq.1)}$$

in which f_n quantifies the contributions of each fragment *n* in a particular chemical (see e.g. Table 1) and F_p accounts for any special intramolecular interaction *p* between the fragments.

This fragment approach has been improved during the last decades and is available in the EPISUITE program from the US Environmental Protection Agency. Other programs for the calculation of K_{ow} values are: ChemProp, and ChemAxon from Chemspider.

Table 1. Fragment constants (K_{ow}) for a few fragments. (from the EPISUITE program)

Fragment	Fragment constant (f)[a]
-CH_3 aliphatic carbon	0.5473
Aromatic Carbon	0.2940
-OH hydroxy, aromatic attach	-0.4802
-N aliphatic N, one aromatic attach	-0.9170

Note: the above calculations are given for non-ionized chemicals. The hydrophobicity of ionic chemicals is also highly affected by the degree of ionization (see below).

K_{ow} values can also be retrieved from databases like echemportal or ECHA and others.

Volatility

Volatility of a chemical from the aqueous phase to air (see Figure 5) is expressed via the Henry's law constant (K_H).

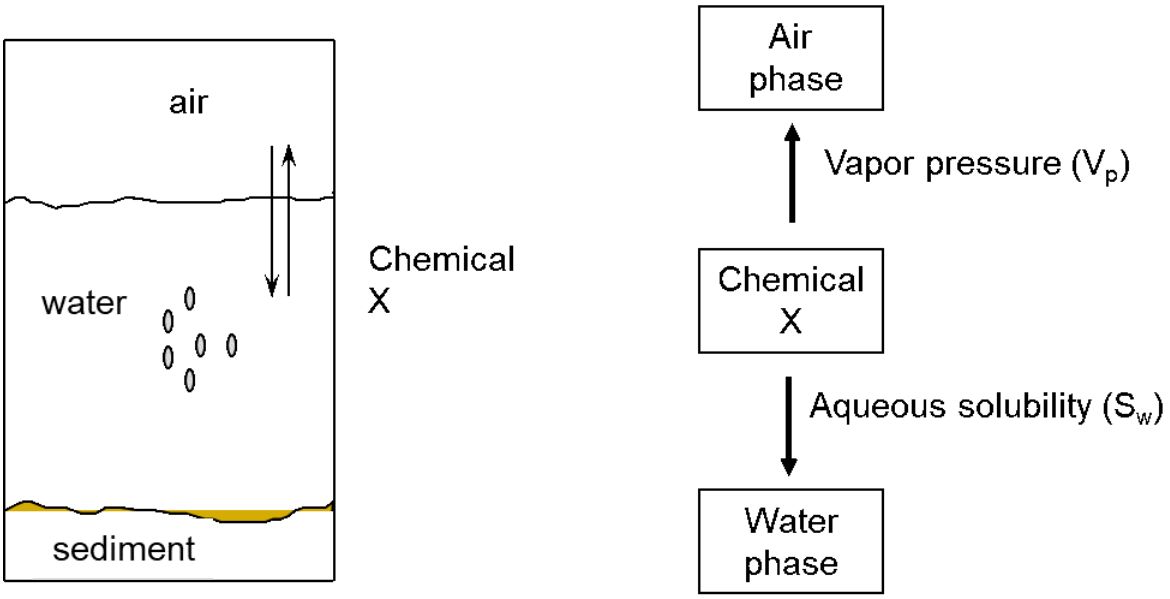


Figure 5.

Evaporation of chemical X from water to air.

Henry's law constant (K_H, in Pa·m^3/mol) is the chemical distribution between the gas phase and water, as

$K_H = \frac{P_i}{C_{aq}}$ (eq.2)

where in an equilibrated water-gas system:

C_{aq} is the aqueous concentration of the chemical (units in mol/m^3), and P_i is the partial pressure of the chemical in air (units in Pascal, Pa), which is the pressure exerted by the chemical in the total gas phase volume (occupied by the mixture of gases the gas-phase above the water solution of the chemical). Note that Pi is a measure of the concentration in the gas phase, but not yet in the same units as the dissolved concentration (discussed below)!

For compounds that are slightly soluble in water, K_H can be estimated from:

$K_H = \frac{V_p}{S_w}$ (eq.3)

where:

K_H: Henry's law constant (Pa·m^3/mol), V_p is the (saturated) vapor pressure (Pa), which is the pressure of the chemical above the pure condensed (liquid) form of the chemical, and S_w is the maximum solubility in water (mol/m^3).

The advantage of equation 3 is that both V_p and S_w can be experimentally derived or estimated. The rationale behind equation 3 is that two opposite forces will affect the evaporation of a chemical from water to air:

(i) the vapor pressure (V_p) of the pure chemical - high vapor pressure means more volatile, and

(ii) solubility in water (S_w) - high solubility means less volatile.

Benzene and ethanol (see Table 2) are good illustrations. Both chemicals have similar vapor pressure, but the Henry's law constant for benzene is much higher because of its much lower solubility in water compared to ethanol; benzene is much more volatile from an aqueous phase.

Table 2. Air-water partition coefficients ($K_{air\text{-}water}$) calculated for five chemicals (ranked by aqueous solubility) by equation 3.

Chemical	Vapor pressure (Pa)	Solubility (mol/m^3)	K_H (Pa.m^3/mol)	$K_{air\text{-}water}$ (L/L, or m^3/m^3)
Ethanol	$7.50 \cdot 10^{3}$	$1.20 \cdot 10^{4}$	$6.25 \cdot 10^{-1}$	$2.53 \cdot 10^{-4}$
Phenol	$5.50 \cdot 10^{1}$	$8.83 \cdot 10^{2}$	$6.23 \cdot 10^{-2}$	$2.52 \cdot 10^{-5}$
Benzene	$1.27 \cdot 10^{4}$	$2.28 \cdot 10^{1}$	$5.57 \cdot 10^{2}$	$2.25 \cdot 10^{-1}$
Pyrene	$6.00 \cdot 10^{-4}$	$6.53 \cdot 10^{-4}$	$9.18 \cdot 10^{1}$	$3.71 \cdot 10^{-4}$
DDT	$2.00 \cdot 10^{-5}$	$2.82 \cdot 10^{-6}$	7.08	$2.86 \cdot 10^{-3}$

Note: all chemicals at equilibrium have a higher concentration (in e.g. mol/L) in the aqueous phase than in the gas phase. Of these five, benzene is the chemical most prone to leave water, with an equilibrated air concentration about 4 times lower (22.5%) than the dissolved concentration.

Equations 2 and 3 are based on the pressure in the gas phase. Environmental fate is often based on partition coefficients, in this case the air-water partition coefficient ($K_{air\text{-}water}$). These partition coefficients are more or less 'dimensionless', because the concentrations are based on equal volumes (such as L/L), while K_H has the unit Pa·m^3/mol or something equivalent to the applied units (equation 4).

$$K_{air-water} = \frac{C_{air}}{C_{aq}} \quad \text{(eq.4)}$$

where:

C_{air} is the concentration in air (in e.g. mol/m^3) and C_{aq} is the aqueous concentration (in e.g. mol/m^3).

$K_{air\text{-}water}$ can be calculated from K_H according to equation 5:

$$K_{air-water} = \frac{K_H}{RT} \quad \text{(eq.5)}$$

where R is the gas constant (8.314 m^3·Pa·K^{-1}·mol^{-1}), and T is the temperature in Kelvin (Kelvin = ^{o}Celsius + 273).

This use of "RT" converts this gas phase concentration to a volume based metric, and applies the ideal gas law which relates pressure (P, in Pa) to temperature (T, in K), volume (V, in m^3), and amount of gas molecules (n, in mol), according to the gas constant (R: 8.314 m^3·Pa·K^{-1}·mol^{-1}):

P·V = n·R·T (note that the units of both terms will cancel out) (eq.6)

At 25 ^{o}Celcius (298 K), the product RT equals 2477 m^3·Pa·K^{-1}·mol^{-1}.

Examples of calculated values for $K_{air\text{-}water}$ are presented in Table 2.

The influence of the chemical structure on volatility of a chemical from a solvent fully depends on the cost of creating a cavity in the solvent (interactions between solvent molecules) and the interactions between the chemical and the solvent molecules. For partitioning processes, the gas phase is mostly regarded as an inert compartment without chemical interactions (i.e. gas phase molecules hardly ever touch each other).

The molecules of a strongly dipolar solvent such as water that contain atoms that can interact as hydrogen acceptor (the O in an OH group) and hydrogen donor (the H in an OH group) strongly interact with each other, and it costs much energy to create a cavity. This cost increases strongly with molecular size, for nearly all molecules more than the energy regained by interactions with the surrounding solvent molecules. As a result, for most classes of organic chemicals, affinity with water decreases and volatility out of water into air slightly increases with molecular volume. For chemicals that are not able to re-interact via hydrogen bonding, e.g. alkanes, the overall volatility is much higher than for chemicals that do have specific interactions with water molecules besides van der Waals.

Degree of ionization

Acids and bases can be present in the neutral (HA and B) or ionized form (A^- and BH^+, respectively). For acids, the neutral form (HA) is in equilibrium with the anionic form (A^-) and for bases the neutral form (B) is in equilibrium with the cationic form (BH^+). The degree of ionization depends on the pH and the acid dissociation constant (pKa). Table 3 shows the equations to calculate the fraction ionized for acids and bases and examples of two acids (phenols) are presented in Table 4.

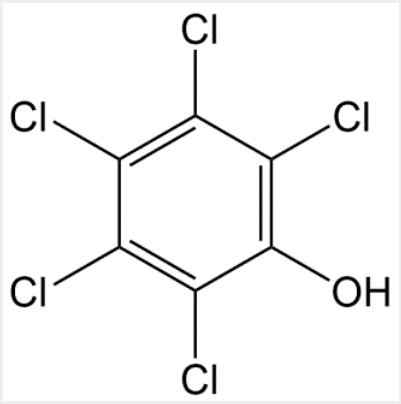
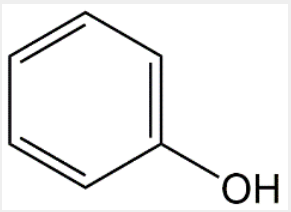
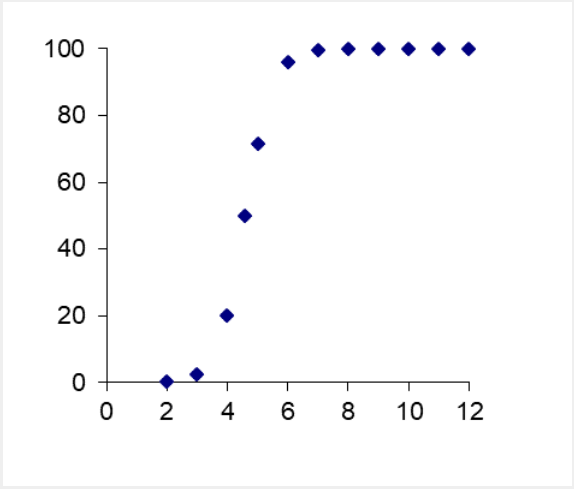
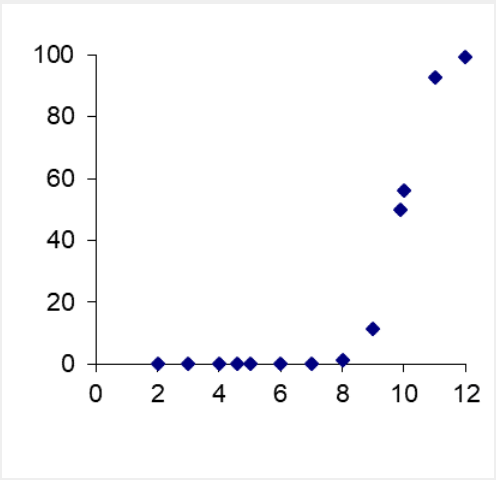
Table 3. Calculation of the fraction ionized for acids and bases.

Acids	**Bases**
	
$fraction\ ionized = \frac{1}{1 + 10^{(pK_a - pH)}}$	$fraction\ ionized = \frac{1}{1 + 10^{(pH - pK_a)}}$

pK_a = - log K_a, where K_a is dissociation constant of the acidic form (HA or BH^+).

The degree of ionization is thus determined by the pH and the pKa value and more examples for several organic chemicals are presented elsewhere (see Chapter Ionogenic organic compounds).

Table 4. The degree of ionization of two phenolic structures (acids).

Pentachlorophenol	**Phenol**
	
pKa = 4.60	pKa = 9.98
% ionized versus pH 	% ionized versus pH 
at pH of 7.0: 99.6 % ionized	at pH of 7.0: 0.1 % ionized

Examples for several organic chemicals are presented elsewhere (see section on Ionogenic organic compounds).

The fate of ionic chemicals is very different from that of non-ionic chemicals. The sediment-water sorption coefficient of the anionic species is substantially (>100x) lower than that of the neutral species. If the percentage of ionization is less than ~99 % (at a pH 2 units above the pKa), the sorption of the anion may be neglected (K_d is still dominated by the >1% neutral species) (Schwarzenbach et al., 2003). The reason for the low sorption affinity of the anionic acid form is twofold: anions are much better water soluble, but also most sediment particles (clay, organic matter, silicates) are negatively charged, and electrostatically repulse the similarly charged chemical. In that case the sorption coefficient K_d can be calculated from the sorption coefficient of the non-ionic form and the fraction of the non-ionized form (α):

$K_d = \alpha \, K_d \, (neutral \; form)$ (eq. 7)

In environments where the pH is such that the neutral acid fraction <1% (when pH >2 units above the pKa), the sorption of the anionic species to soil/sediment may significantly contribute to the overall "distribution coefficient" of both acid species.

For basic environmental chemicals of concern, among which many illicit drugs (e.g. amphetamine, cocaine) and non-illicit drugs (e.g. most anti-depressants, beta-blockers), the protonated forms are positively charged. These organic cations are also much more soluble in water than the neutral form, but at the same time they are electrostatically attracted to the negatively charged sediment

surfaces. As a result, the sorption affinity of organic cations to sediment should not be considered negligible relative to the neutral species. The sorption processes, however, may strongly differ for the neutral base species and the cationic base species. Several studies have shown that the sorption affinity of cationic base species to DOM or sediment is even stronger than that of the neutral species.

References

De Bruijn, J., Busser, F., Seinen, W., Hermens, J. (1989). Determination of octanol/water partition coefficients for hydrophobic organic chemicals with the "slow-stirring" method. Environmental Toxicology and Chemistry 8, 499-512.

Eadsforth, C.V. (1986). Application of reverse-phase HPLC for the determination of partition coefficients. Pesticide Science 17, 311-325.

Leo, A., Hansch, C., Elkins, D. (1971). Partition coefficients and their uses. Chemical Reviews 71, 525-616.

Leo, A.J. (1993). Calculating log P(oct) from structures. Chemical Reviews 93, 1281-1306.

Rekker, R.F., de Kort H.M. (1979). The hydrophobic fragmental constant; an extension to a 1000 data point set. *Eur.J. Med. Chem. - Chim. Ther.* 14:479-488.

Schwarzenbach RP, Gschwend PM, Imboden DM (Eds.) 2003. Environmental organic chemistry. Wiley, New York, NY, USA.

Further reading:

Mackay, D., Boethling, R.S. (Eds.) 2000. Handbook of property estimation methods for chemicals: environmental health and sciences. CRC Press.

van Leeuwen, C.J., Vermeire, T.G. (Eds.) 2007. Risk assessment of chemicals: An introduction. Springer, Dordrecht, The Netherlands

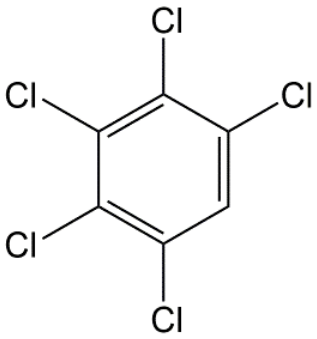
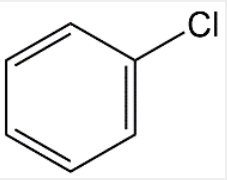
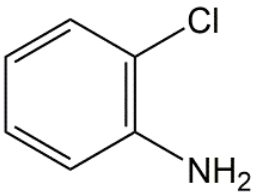
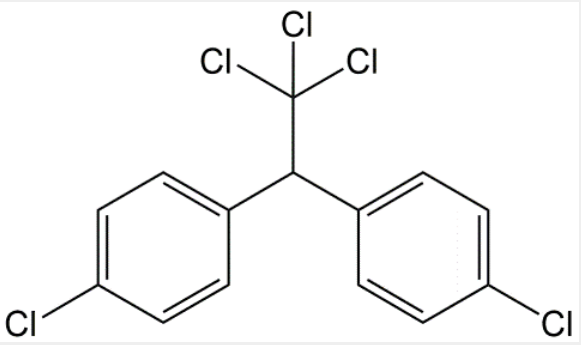
3.4.1. Question 1

Explain the term hydrophobicity and mention two major properties that affect hydrophobicity of chemicals.

3.4.1. Question 2

What is the definition of K_{ow}. Rank the following chemicals from the Table below from low to high K_{ow}. 1. Pentachlorobenzene, 2.

Monochlorobenzene, 3. Monochloroaniline, 4. DDT.

1.Pentachlorobenzene	
2.Monochlorobenzene	
3.Monochloroaniline	
4.DDT	

3.4.1. Question 3

Rank the chemicals in the Table according to their volatility as pure compounds and their solubility in water and volatility from water.

3.4.1. Question 4

Which two basic properties determine the volatility of a chemical from water to air.

3.4.1. Question 5

Calculate the percentage ionized for 2,3,4-trichlorophenol (pKa = 4.6) at pH 3, 5, 7 and 9.

3.4.2. Sorption

Authors: Joop Hermens

Reviewers: Kees van Gestel, Steven Droge, Philipp Mayer

Leaning objectives:

You should be able to:

- understand why information on sorption is important for risk assessment
- give examples that illustrate the importance of sorption for risk assessment
- understand the concept of sorption isotherms
- be familiar with different sorption isotherms (linear, Freundlich, Langmuir).

Keywords: Sorption isotherm, absorption and adsorption, organic matter, Freundlich model, Langmuir model, organic carbon content

Introduction

Sorption processes have a major influence on the fate of chemicals in the environment (Box 1). In general, sorption is defined as the binding of a dissolved or gaseous chemical (the sorbate) to a solid phase (the sorbent) and this may involve different processes, including:

(i) binding of dissolved chemicals from water to sediments and soils

and (ii) binding of gaseous phase chemicals from air to soils, plants, and trees.

Information about sorption is relevant because of a number of reasons:

- sorption controls the actual fate and thereby the risk of (many) organic and inorganic contaminants in the environment,
- sorbed chemicals cannot evaporate, are not available for (photo)chemical or microbial breakdown, are not as easily transported as dissolved/vapor phase chemicals, and are not available for uptake by organisms,
- sorption also plays an important role in toxicity tests, affecting exposure concentrations.

Box 1.
The Biesbosch is a wetland area in the Netherlands, an area in between the Rivers Rhine and Meuse and estuaries that are connected to the North Sea. The water flow is relatively low and as a consequence there is a strong sedimentation of particles from the water to the sediment. Chemicals present in the water strongly sorb to these particles, which in the past were polluted with hydrophobic organic contaminants such as dioxins and PCBs. The concentrations of these organic compounds in sediment are still relatively high because they are highly persistent. The reason for this persistence of these compounds is that these sorbed compounds are not easily available for degradation by bacteria. Also, the concentrations in organisms that live close to or in the sediment are high. These concentrations are so high that fishing on eel, for example, is not allowed in the area. This example shows the importance of sorption processes on fate, but also on effects in the environment.

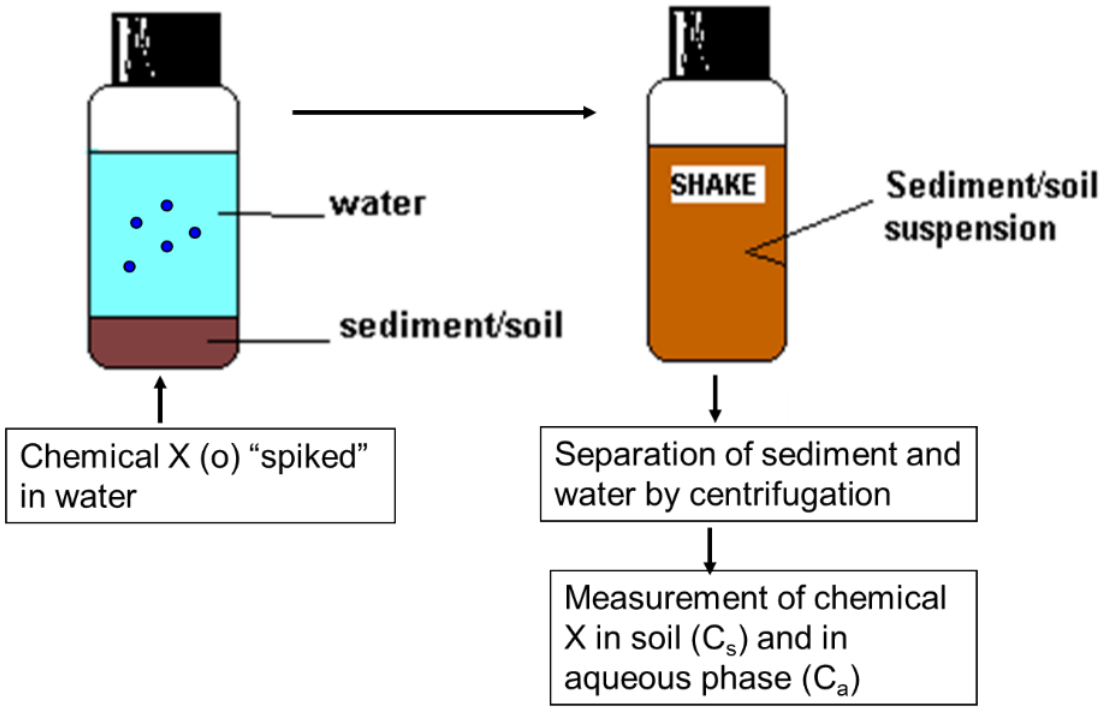


Figure 1. Measurement of sorption coefficients.

Measurement of sorption is a simple procedure. A chemical X is spiked (added) to the aqueous phase in the presence of a certain amount of the solid phase (sediment or soil). The chemical sorbs to the solid phase and when the system is in equilibrium, the concentrations in the sediment (C_s) and in the aqueous phase (C_a) are measured. The solid phase is collected via centrifugation or filtration.

The sorption coefficient K_p (equation 1 and box 2) gives information about the degree of sorption of a chemical to sediment and is defined as:

$$K_p = \frac{C_s}{C_a}$$ (1)

Box 2:

The concentration of a chemical X in sediment (C_s) is 30 mg/kg and the concentration in the aqueous phase (C_a) is 0.1 mg/L.

The sorption coefficient $K_p = C_s / C_a = 30$ mg/kg / 0.1 mg/L = 300 L/kg

Note the units of a sorption coefficient: L/kg

In the environmental risk assessment of chemicals, it is very useful to understand the fraction of the total amount of chemical (A_{total}) in a system that is sorbed (f_{sorbed}) or dissolved ($f_{dissolved}$) (e.g. due to an accidental spill in a river):

$f_{dissolved} = A_{dissolved} / A_{total}$, and thus $f_{sorbed} = 1 - f_{dissolved}$

This is related to the sorption coefficients of X and the volume of the solvent and the volume of the sorbent material. The equation derived for calculating $f_{dissolved}$ is based on the mass balance of chemical A, which relates the concentration of X (C) to the amount of X (A) in each volume (V):

$C = A / V$, and thus $A = C \cdot V$

which for a system of water and sediment (air not included for simplification) relates to:

$A_{total} = A_{dissolved} + A_{sorbed} = C_{water} \cdot V_{water} + C_{sediment} \cdot V_{sediment} = C_{water} \cdot V_{water} + (K_p \cdot C_{water}) \cdot V_{sediment}$

$f_{dissolved} = A_{dissolved} / A_{total} = C_{water} \cdot V_{water} / (C_{water} \cdot V_{water} + K_p \cdot C_{water} \cdot V_{sediment})$

This way of separating out $C_{sediment}$ from the equation using K_p can result, after rearranging (by dividing both parts of the ratio by $C_{water} \cdot V_{water}$) to the following simplified equation:

$f_{dissolved} = 1 / (1 + K_p \cdot (V_{sediment} / V_{water}))$

in this equation, 'sediment' can be replaced by any sorbent, as long as the appropriate sorption coefficient is used.

Let's try to calculate with chemical X from above, in a wet sediment, where 1L wet sediment contains ~80% water and 20% dry solids. The dissolved fraction of X with K_p = 300 kg/L, is only 0.013 in this example. Thus, with 1.3% of X actually dissolved, this indicates that 98.7% of X is sorbed to the sediment.

Sorption processes

There are two major sorption processes (see Figure 2):

- Absorption - partitioning ("dissolution") of a chemical into a 3-D sorbent matrix. The concentration in the sorbing phase is homogeneous.
- Adsorption - binding of a chemical to a 2-D sorbent surface. Because the number of sorption sites on a surface is limited, sorption levels off at high concentrations in the aqueous phase.

A sorption isotherm gives the relation between the concentration in a sorbent (sediment) and the concentration in the aqueous phase and the isotherm is important in identifying a sorption process.

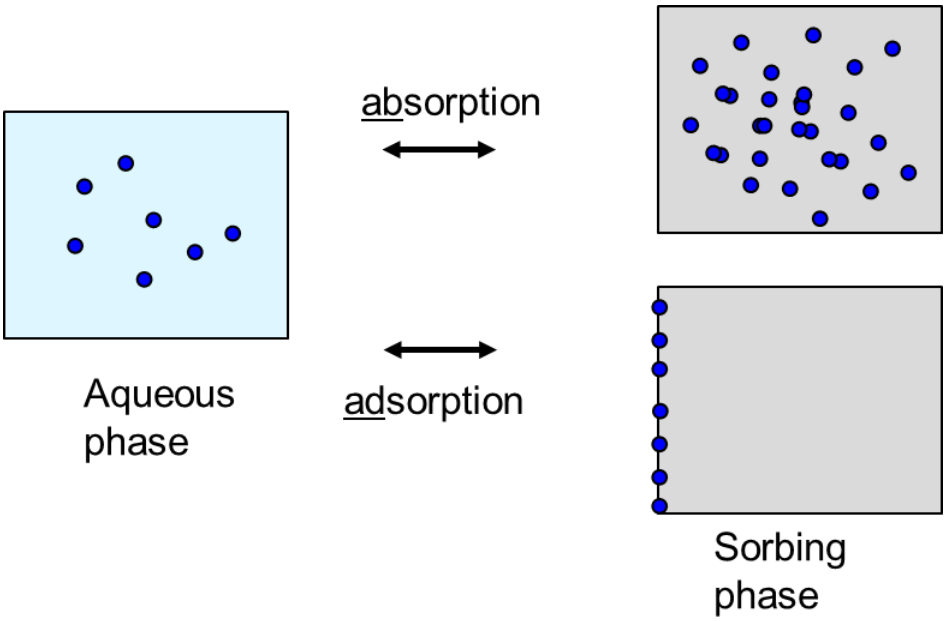


Figure 2. Two sorption processes: absorption and adsorption.

Absorption of a chemical is similar to its partitioning between two phases and comparable to its partitioning between two solvents. Distribution of a chemical between octanol and water is a well-known example of a partitioning process (see Section 3.4.1 on Relevant chemical properties for more detailed information on octanol-water partitioning). The isotherm for an absorption process is linear (Figure 3A) and the slope of the y-x plot is the sorption coefficient K_p.

A. Linear

K_p is the sorption coefficient

$$K_p = \frac{C_s}{C_{aq}}$$

$$\text{Log } C_s = \text{Log } (C_{aq}) + \text{Log } (K_p)$$

B. Langmuir

b is a constant related to the binding energy

C_{max} is the maximum sorption capacity

$$C_s = \frac{b \,.\, C_{max} \,.\, C_{aq}}{1 + b \,.\, C_{max}}$$

C. Freundlich

K_F is the unit-dependent Freundlich coefficient

n is the Freundlich exponent

$$C_s = K_F \cdot C_{aq}^{\,n}$$

$$\text{Log } C_s = n \cdot \text{Log } (C_{aq}) + \text{Log } (K_F)$$

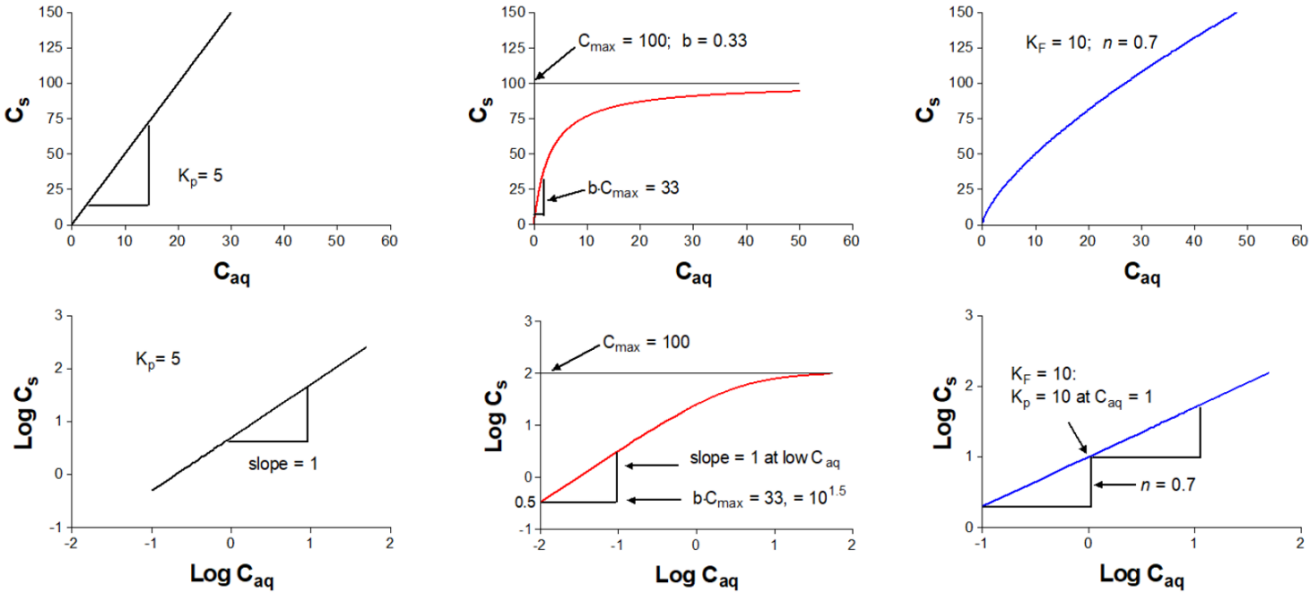


Figure 3. Sorption isotherms for absorption: linear model (3A), and for adsorption: Langmuir model (3B) or Freundlich model (3C).

In an adsorption process, where the sorbing phase is a surface with a limited number of sorption sites, the sorption isotherm is non-linear and may reach a maximum concentration that is adsorbed when all sites are occupied. A mechanistic model for adsorption is the Langmuir model. This model describes adsorption of molecules to homogeneous surfaces with equal adsorption energies, represented by the adsorption site energy term (b) and a limited number of sorption sites (C_{max}) that can become saturated (Figure 3B). The Langmuir adsorption coefficient (K_{ad}) is equal to the product ($b \cdot C_{max}$) at relatively low aqueous concentrations, where the product ($b \cdot C_{aq}$) << 1 (note that the denominator term will then be ~1). Indeed, the isotherm curve on a double log scale plot shows a slope of 1 at such low concentrations, indicating linearity.

Another mathematical approach to describe non-linear sorption is the Freundlich isotherm (Figure 3C), where K_F is the Freundlich sorption constant and n is the Freundlich exponent describing the sorption process non-linearity. Using logarithmic values for aqueous and sorbed concentrations, the Freundlich isotherm can be rewritten as:

$\text{Log } C_s = n \cdot \log C_{aq} + \log K_F$ (eq. 2)

This conveniently yields a linear relationship (just as y = a·x + b) between log Cs and log Caq, with a slope equal to n and the abscissa (crossing point with the Y-axis) equal to log K_F. This allows for easy fitting of linear trend lines through experimental data sets. When n = 1, the isotherm is linear, and equals the one for absorption. In case of saturation of the sorption sites on the solid phase, 1/n will be smaller than 1. The Freundlich isotherm can, however, also yield a 1/n value > 1; this may occur for example if the chemical that is sorbed itself forms a layer that serves as a new sorbing phase and examples are described for surfactants.

https://bio.libretexts.org/@go/page/108675

Sorption phases

Soils and sediments may show large variations in composition and particle size distribution. The major components of soils and sediments are:

Sand	63 - 2 mm
Silt	2 - 63 µm
Clay	<2 µm
Organic matter	includes e.g. detritus, humic acids, especially associated with the clay and silt fractions
$CaCO_3$	

Figure 4 gives a schematic picture of a sediment or soil particle. In addition to the presence of clay minerals and (soil or sediment) organic matter (SOM), sediment and soil may contain soot particles (a combustion residue).

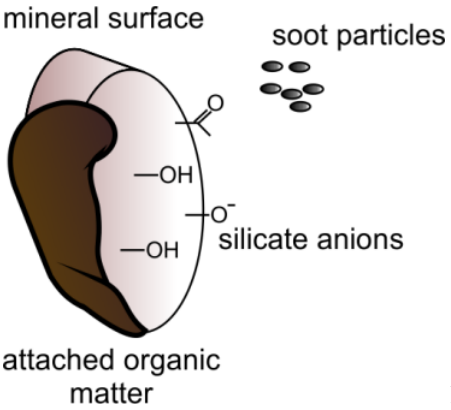


Figure 4. Structure of a soil or sediment particle showing the major components: organic matter and clay mineral and soot. Modified from Schwarzenbach et al. (2003) by Steven Droge.

Organic matter is formed upon decomposition of plant material and dead animal or microbial tissues. Upon decomposition of plant material, the first organic groups to be released are phenolic acids, some of which have a high affinity for complexation of metals. One example is salicylic acid (o-hydroxybenzoic acid), which occurs in high concentrations in leaves of willows, poplar and other deciduous trees. Further decomposition of plant material may result in the formation of humic acids, fulvic acids and humin. Humic and fulvic acids contain a series of functional groups, such as carboxyl- (COOH), carbonyl- (=C=O), phenolic hydroxyl- (-OH), methoxy- (-OCH3), amino- (-NH2), imino (=NH) and sulfhydryl (-SH) groups (see for more details the section on Soil).

Hydrophobic organic chemicals mainly sorb to organic matter. Because organic matter has the characteristics of a solvent, the sorption is clearly an absorption process and the sorption isotherm is linear. Because binding is mainly to organic matter, the sorption coefficient (K_p) depends on the fraction of organic matter (f_{om}) or the fraction of organic carbon (f_{oc}) present in the soil or sediment. Please note that as a rule of thumb, organic matter contains 58% organic carbon ($f_{oc} = 0.58 \cdot f_{om}$). Figure 5A shows the increase in sorption coefficient with increasing fraction organic carbon in soils and sediments. In order to arrive at a more intrinsic parameter, sorption coefficients are often normalized to the fraction organic matter (K_{om}) or organic carbon (K_{oc}). These K_{oc} or K_{om} values are less dependent of the sediment or soil type (Figure 5B).

$$K_{om} = \frac{K_p}{f_{om}} \quad (3)$$

$$K_{oc} = \frac{K_p}{f_{oc}} \quad (4)$$

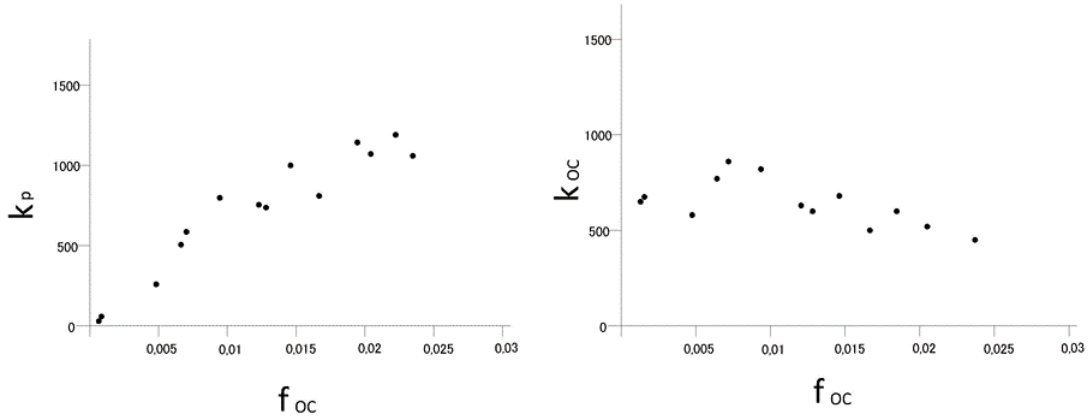


Figure 5. The relationship between the sorption coefficient (K_p) (left) and the organic carbon normalized sorption coefficient (K_{oc}) (right) and the fraction organic carbon (f_{oc}). Data from Means et al. (1980). Drawn by Wilma Ijzerman.

Hydrophobic chemicals can have a very high affinity to soot particles relative to the affinity to SOM. If a sediment contains soot, K_p values are often higher than predicted based on the fraction organic carbon in the organic matter (Jonker and Koelmans, 2002).

References

Schwarzenbach, R.P., Gschwend, P.M., Imboden, D.M. (2003). Environmental Organic Chemistry. Wiley, New York, NY, USA.

Means, J.C., Wood, S.G., Hassett, J.J., Banwart, W.L. (1980). Sorption of polynuclear aromatic-hydrocarbons by sediments and soils. Environmental Science and Technology 14, 1524-1528.

Jonker, M.T.O., Koelmans, A.A. (2002). Sorption of polycyclic aromatic hydrocarbons and polychlorinated biphenyls to soot and soot-like materials in the aqueous environment mechanistic considerations. Environmental Science and Technology 36, 3725-3734.

Suggested reading

van Leeuwen, C.J., Vermeire, T.G. (Eds.) (2007). Risk Assessment of Chemicals: An Introduction. Springer, Dordrecht, The Netherlands. Chapter 3 and 9.

Schwarzenbach, R.P., Gschwend, P.M., Imboden, D.M. (2003). Environmental Organic Chemistry. Wiley, New York, NY, USA. chapters 9, 11. Detailed information about sorption processes and sorption mechanisms.

3.4.2. Question 1

What is a sorption isotherm and what is the difference between absorption and adsorption?

3.4.2. Question 2

Why is the sorption isotherm to clay non-linear?

3.4.2. Question 3

What are the main sorption phases in sediment or soil?

3.4.3. QSPRs

Author: Joop Hermens

Reviewer: Steven Droge, Monika Nendza

Learning objectives:

You should be able to:

- indicate which properties of chemicals are applied in a QSPR
- list different techniques to derive a QSPR
- indicate which interactions may occur between molecules.

Keywords: Quantitative structure-property relationships (QSPR), quantitative structure-activity relationships (QSAR), octanol-water partition coefficients, hydrogen bonding, multivariate techniques.

Introduction

Risk assessment needs input data for fate and effect parameters. These data are not available for many of the existing chemicals and predictions via estimation models will provide a good alternative to actual testing. Examples of estimation models are Quantitative Structure-Property Relationships (QSPRs) and Quantitative Structure-Activity Relationships (QSARs). The term "activity" is often used in relation to models for toxicity, while "property" usually refers to physical-chemical properties or fate parameters.

In a QSAR or QSPR, a certain environmental parameter is related to a physical-chemical or structural property, or a combination of properties.

The elements in a QSPR or QSAR are shown in Figure 1 and include:

- the parameter for which the estimation model has been developed: the Y-variable (upper right),
- the properties of the chemical or chemical parameter: the X-variable (upper left),
- the model itself (center), and
- the prediction of a fate or effect parameter from the chemical properties (bottom).

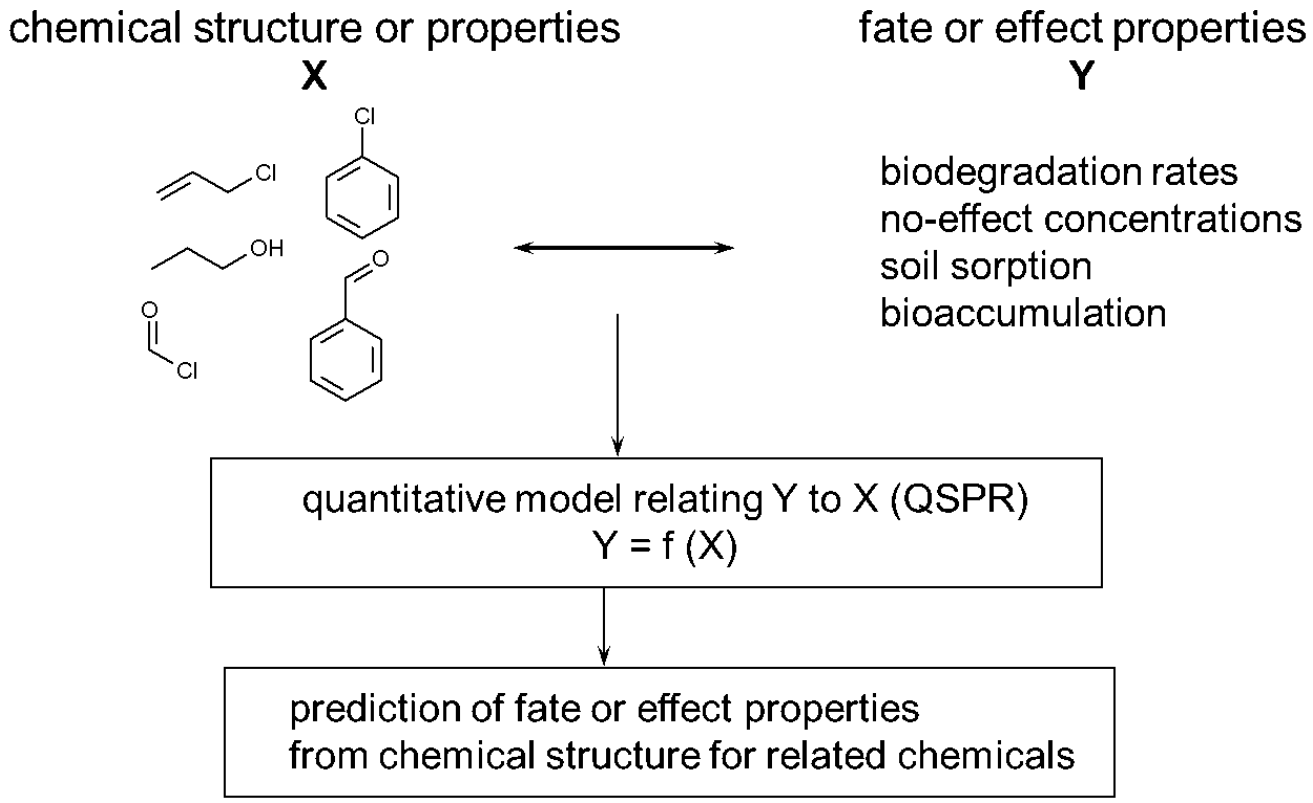


Figure 1. The principle of a QSPR or QSAR. See text for explanation.

The Y-variable

Estimation models have been developed for many endpoints such as sorption to sediment, humic acids, lipids and proteins, chemical degradation, biodegradation, bioconcentration and ecotoxic effects.

The X-variable

An overview of the chemical parameters (the X-variable) used in estimation models is given in Table 1. Chemical properties are divided in three categories: (i) parameters related to hydrophobicity, (ii) parameters related to charge and charge distribution in a

molecule and (iii) parameters related to the size or volume of a molecule. Hydrophobicity is discussed in more detail in the section on Relevant chemical properties.

Other QSPR approaches use large number of parameters derived from chemical graphs. The CODESSA Pro software, for example, generates molecular (494) and fragment (944) descriptors, classified as (i) constitutional, (ii) topological, (iii) geometrical, (iv) charge related, and (v) quantum chemical (Katritzky et al. 2009). Some models are based on structural fragment in a molecule. The polyparameter linear free energy relationships (pp-LFER) use parameters that represent interactions between molecules (see under pp-LFER).

Table 1. Examples of parameters related to hydrophobicity and electronic and steric parameters (the X variable).

Hydrophobic parameters
Aqueous solubility
Octanol-water partition coefficient (K_{ow})
Hydrophobic fragment constant π
Electronic parameters
Atomic charges (q)
Dipole moment
Hydrogen bond acidity (H bond-donating)
Hydrogen bond basicity (H bond-accepting)
Hammett constant σ
Steric parameters
Total Surface Area (TSA)
Total Molecular Volume (TMV)
Taft constant for steric effects (Es)

The model

Most models are based on correlations between Y and X. Such a relationship is derived for a "training set" that consists of a limited number of carefully selected chemicals. The validity of such a model should be tested by applying it to a "**validation set**", i.e. a set of compounds for which experimental data can be compared with the predictions. Different techniques can be used to develop an empirical model, such as:

- graphical presentations,
- linear or non-linear equations between Y and X,
- linear or non-linear equations based on different properties (Y versus X_1, X_2, etc.),
- multivariate techniques such as Principal Component Analysis (PCA) and Partial Least Square Analysis (PLS).

Linear equations take the form:

$$Y(i) = a_1X_1(i) + a_2X_2(i) + a_3X_3(i) + ... + b \quad (1)$$

where Y(i) is the value of the dependent parameter of chemical i (for example sorption coefficients); X_1-X_3(i) are values for the independent parameters (the chemical properties) of chemical i; a_1-a_3 are regression coefficients (usually 95% confidence limits are given); b is the intercept of the linear equation. The quality of the equation is presented via the correlation coefficient (r) and the standard error of estimate (s). The closer r is to 1.0, the better the fit of the relationship is. More information about the statistical quality of models can be found under "limitation of QSPR".

The classical approach in QSAR and QSPR studies is the Hansch approach that was develop in the 1960s. The Hansch equation (Hansch et al., 1963) describes the influence of substituents on the biological activity in a series of parent compounds with a certain

substituent (equation 2). Substituents are for example a certain atom or chemical group (Cl, F, B,. OH, NH_2) attached to a parent aromatic ring structure.

$\log 1/C = c\,\pi + c'\,\sigma + c''\,E_s + c'''$ (2)

in which:

C is the molar concentration of a chemical with a particular effect,

π is a substituent constant for hydrophobic effects,

σ is a substituent constant for electronic effects, and

E_s is a substituent constant for steric effects.

c are constants that are obtained by fitting experimental data

For example, the hydrophobic substituent constant is based on K_{ow} and is defined as is defined as:

$\pi\,(X) = \log K_{ow}\,(RX) - \log K_{ow}\,(RH)$ (3)

where RX and RH are the substituted and unsubstituted parent compound, respectively.

The Hammett and Taft constants are derived in a similar way.

Multivariate techniques may be very useful to develop structure-activity relationships, in particular in cases where a large number of chemical parameters is involved. Principal Component Analysis (PCA) can be applied to reduce the number of variables into a few principal components. The next step is to find a relationship between Y and X via, for example, Partial Least Square (PLS) analysis. The advantage of PCA and PLS is that it can deal with a large number of chemical descriptors and that is can also cope with collinear (correlated) properties. More information on these multivariate techniques and examples in the field of environmental science are given by Eriksson et al. (1995).

Poly-parameter Linear Free Energy Relationship (pp-LFER)

The pp-LFER approach has a strong mechanistic basis because it includes the different types of interactions between molecules (Goss and Schwarzenbach, 2001). For example, the sorption coefficient of a chemical from an aqueous phase to soil or to phospholipids (the sorbent) depends on the interaction of a chemical with water and the interaction with the sorbent phase. One of the driving forces behind sorption is the hydrophobicity. Hydrophobicity means fear (phobia) for water (hydro). A hydrophobic chemical prefers to "escape from the aqueous phase" or in other words "it does not like to dissolve in water". Water molecules are tightly bound to each other via hydrogen bonds. For a chemical to dissolve, a cavity should be formed in the aqueous phase (Figure 2) and this will cost energy. More hydrophobic compounds will often have a stronger sorption (see more information in the section on Relevant chemical properties).

Hydrophobicity mainly depends on two molecular properties:

- Molecular size
- Polarity / ability to interact with water molecules, for example via hydrogen bonding

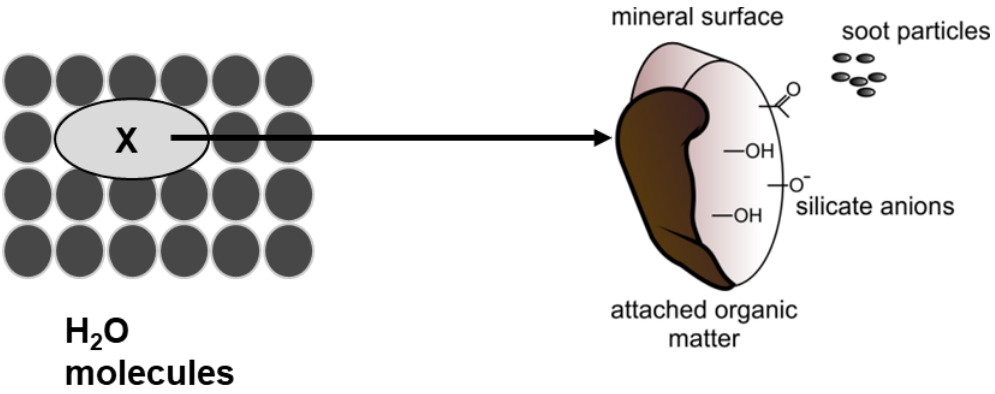


Figure 2. The formation of a cavity in water for chemical X and the interaction with another phase (here, a soil particle).

In the interaction with the sorbent (soil, membrane lipids, storage lipids, humic acids), major interactions are van der Waals interactions and hydrogen bonding (Table 2). Van der Waals interactions are attractive and occur between all kind of molecules and the strength depends on the contact area. Therefore, the strength of van der Waals interactions are related to the size of a molecule. A hydrogen bond is an electrostatic attraction between a hydrogen (H) and another electronegative atom bearing a lone pair of electrons. The hydrogen atom is usually covalently bound to a more electronegative atom (N, O, F). Table 2 lists the interactions with examples of chemical structures.

A pp-LFER is a linear equation developed to model partition or sorption coefficients (K) using parameters that represent the interactions (Abraham, 1993). The model equation is based on five descriptors:

$$logK = c + e \bullet E + s \bullet S + a \bullet A + b \bullet B + v \bullet V \quad (2)$$

with:

E	excess molar refraction
S	dipolarity/polarizability parameter
A	solute H-bond acidity (H-bond donor)
B	solute H-bond basicity (H-bond acceptor)
V	molar volume

The partition or sorption coefficient K may be expressed as the sum of five interaction terms, with the uppercase parameters describing compound specific properties. E depends on the valence electronic structure, S represents polarity and polarizability, A is the hydrogen bond (HB) donor strength (HB acidity), B the HB acceptor strength (HB basicity), V is the so-called characteristic volume related to the molecule size, and c is a constant. The lower-case parameters express the corresponding properties of the respective two-phase system, and can thus be taken as the relative importance of the compound properties for the particular partitioning or sorption process. In this introductory section, we only focus on the volume factor (V) and the two hydrogen bond parameters (A and B).

Numerous pp-LFERs have been developed for all kinds of environmental processes and an overview is given by Endo and Goss (2014).

Table 2. Types of interactions between molecules and the phase to which they sorb with examples of chemicals (Goss and Schwarzenbach, 2003).

Compound[a)]	**Interactions**	**Examples**
Apolar	only van der Waals	alkanes, chlorobenzenes, PCBs
Monopolar	van der Waals + H-acceptor (e-donor)	alkenes, alkynes, alkylaromatic compounds ethers, ketones, esters, aldehydes
Monopolar	van der Waals + H-donor (e-acceptor)	$CHCl_3$, CH_2Cl_2
Bipolar	van der Waals + H-donor + H-acceptor	$R\text{-}NH_2$, R2-NH, R-COOH, R-OH

[a)] Apolar: no polar group present; mono/dipolar: one or two polar groups present in a molecule

Examples of QSPR for bioconcentration to fish

K_{ow} based model

Predictive models for bioconcentration have a long history. The octanol-water partition coefficient (K_{OW}) is a good measure for hydrophobicity and bioconcentration factors (BCF's) are often correlated to K_{ow} (see more information in section on Bioaccumulation). The success of these K_{OW} based models was explained by the resemblance of partitioning in octanol and bulk lipid in the organisms, at least for neutral hydrophobic compounds. A well-known example of a linear QSAR model for the log BCF (Y variable) based on the log K_{OW} (X variable) (Veith et al., 1979):

log BCF = 0.85 log K_{OW} - 0.70 (5)

Figure 3 gives a classical example of such a correlation for BCF to guppy of a series of chlorinated benzenes and polychlorinated biphenyls. When lipophilic chemicals are metabolised, the relation shown in Figure 3 is no longer valid and BCF will be lower than predicted based on K_{OW}. Another deviation of this BCF-K_{ow} relation can be found for highly lipophilic chemicals with log K_{ow}>7. For such chemicals, BCF often decrease again with increasing Kow (see Figure 3). The apparent BCF curve with Kow as the X variable tends to follow a nonlinear curve with an optimum at log K_{ow} 7-8. This phenomenon may be explained from molecular size: molecules of chemicals like decachlorobiphenyl may be so large that they have difficulties in passing membranes. A more likely explanation, however, is that for highly lipophilic chemicals aqueous concentrations may be overestimated. It is not easy to separate chemicals bound to particles from the aqueous phase (see box 1 in the section on Sorption) and this may lead to measured concentrations that are higher than the bioavailable (freely dissolved) concentration (Jonker and van der Heijden 2007; Kraaij et al. 2003). For example, at a dissolved organic carbon (DOC) concentration of 1 mg-DOC/L, a chemical with a log Koc of 7 will be 90% bound to particles, and this bound fraction is not part of the dissolved concentration that equilibrates with the (fish) tissue. This shows that these models are also interesting because they may show trends in the data that may lead to a better understanding of processes.

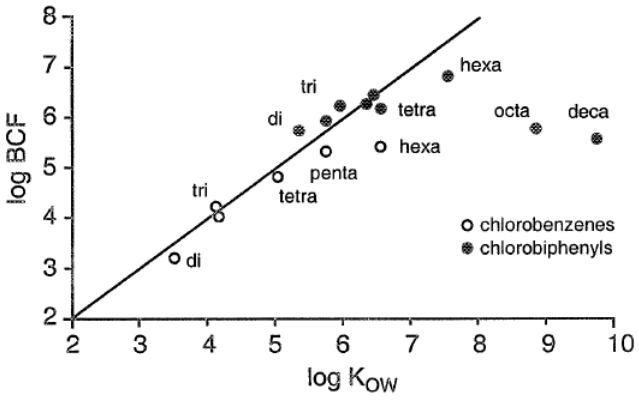
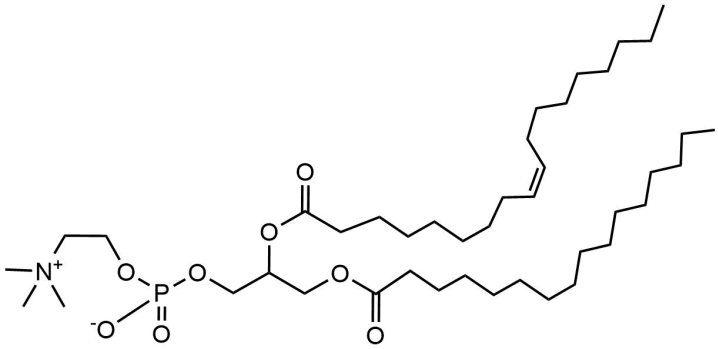


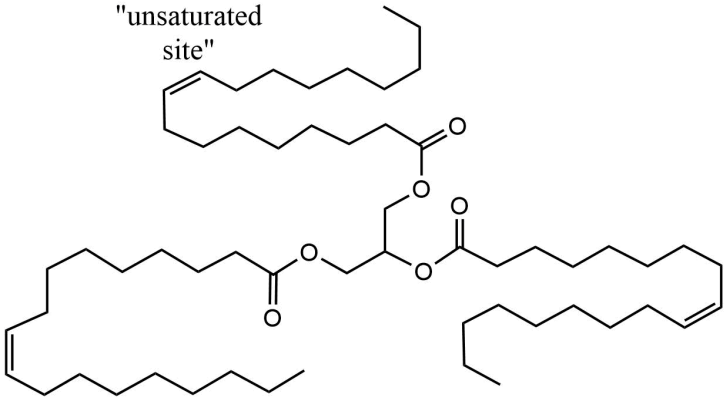
Figure 3. The relationship between bioconcentration factors in guppy and the octanol-water partition coefficients with data from (Bruggeman et al., 1984; Könemann and Leeuwen, 1980).

Examples of QSPR for sorption to lipids

K_{ow} based models are successful because octanol probably has similar properties than fish lipids. There are several types of lipids and membrane lipids have different properties and structure than for example storage lipids (see Figure 4, and more details in the section on Biota). More refined BCF models include separation of storage and membrane lipids and also proteins as separate sorptive phases (Armitage et al. 2013). pp-LFER is a very suitable approach to model these sorption or partitioning processes and results for two large data sets are presented in Table 3. The coefficients *e*, *s*, *b* and *v* are rather similar. The only parameter that is different in these two models is coefficient *a*, which represents the contribution of hydrogen bond (HB) donating properties (A) of chemicals in the data set. This effect makes sense because the phosphate group in the phospholipid structure has strong HB accepting properties. This example shows the strength of the pp-LFER approach because it closely represents the mechanism of interactions.



1-palmitoyl-2-oleoylphosphatidylcholine (POPC), a key phospholipid in cell membranes



Triglyceride, storage fat

Figure 4. Structure of a phospholipid and a triglyceride. Note the similar glycerol part in both lipids.

Table 3. LFERs for storage lipid-water partition coefficients (K_{SL-W}) and membrane lipid-water partition coefficients (K_{ML-W} (liposome)). Listed are the parameters (and standard error), the number of compounds with which the LFER was calibrated (n), the correlation coefficient (r^2), and the standard error of estimate (SE). log K = *c* + *e*E + *s*S + *a*A + *b*B + *v*V.

Para-meter	*c*	*e*	*s*	*a*	*b*	*v*	n	r^2	SE
K_{SL-W}	-0.07 (0.07)	0.70 (0.06)	-1.08 (0.08)	-1.72 (0.13)	-4.14 (0.09)	4.11 (0.06)	247	0.997	0.29
From (Geisler et al. 2012)									
K_{ML-W} (liposome)	0.26 (0.08)	0.85 (0.05)	-0.75 (0.08)	0.29 (0.09)	-3.84 (0.10)	3.35 (0.09)	131	0.979	0.28
From (Endo et al. 2011)									

K_{SL-W}: storage lipid partition coefficients are mean values for different types of oil. Raw data and pp-LFER (for 37 ^{o}C) reported in (Geisler et al. 2012).

K_{ML-W} (liposome): data from liposomes made up of phosphatidylcholine (PC) or PC mixed with other membrane lipids. Raw data (20-40 ^{o}C) and pp-LFER reported in (Endo et al. 2011).

Examples of QSPR for sorption to soil

Numerous QSPRs are available for soil sorption (see section on Sorption). Also the organic carbon normalized sorption coefficient (K_{oc}) is linearly related to the octanol-water partition coefficient (see Figure 5).

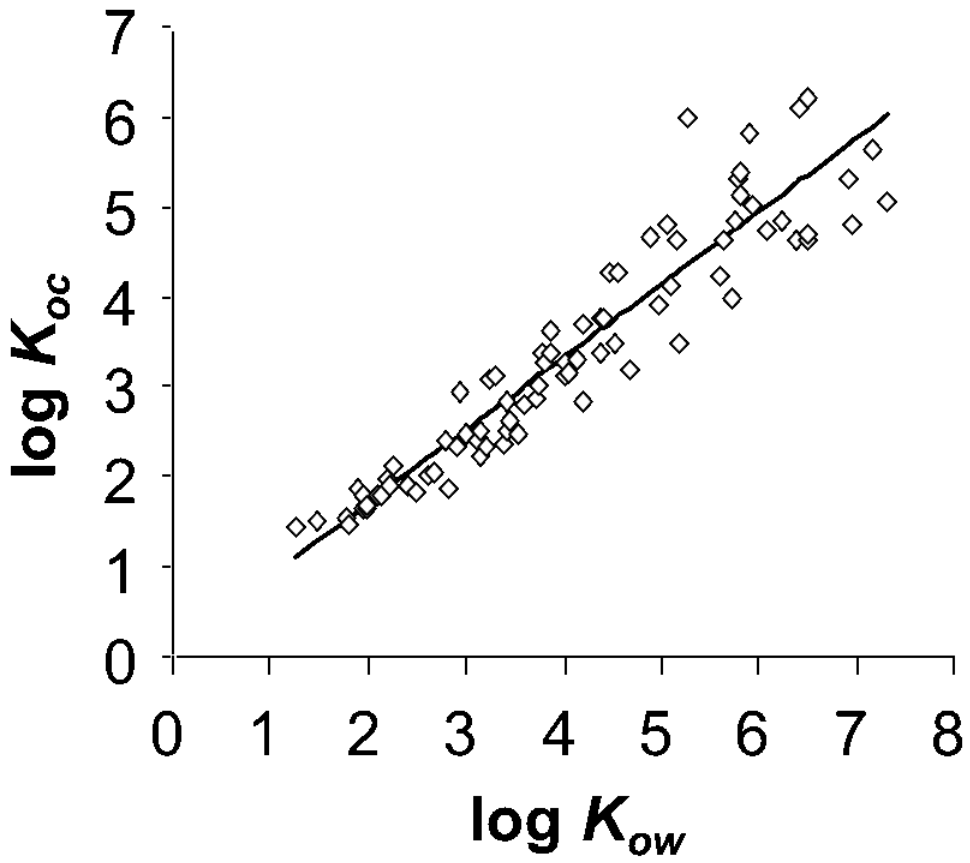


Figure 5. Correlation between the organic carbon normalized sorption coefficient to soil (K_{oc}) and the octanol-water partition coefficient (K_{ow}) for data from (Sabljic et al. 1995).

The model in Figure 5 is only valid for neutral, non-polar hydrophobic organic chemicals such as chlorinated aromatic compounds, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyl (PCBs) and chlorinated insecticides or, in general, compounds that only contain carbon, hydrogen and halogen atoms. It does not apply to polar and ionized organic compounds nor to metals. For polar chemicals, also other interactions may influence sorption and a pp-LFER approach would also be useful.

The sorption of ionic chemicals is more complex. For the sorption of cationic organic compounds, clay minerals can be an equally important sorption phase as organic matter because of their negative surface charge and large surface area. The sorption of organic cations is mainly an adsorption process that reaches a maximum at the cation exchange capacity (CEC) of a particle (see section on Soil). Also models for the prediction of sorption of cationic compounds are more complicated and first attempts have been made recently (Droge and Goss, 2013). Major sorption mechanism for anionic chemicals is sorption into organic matter. The sorption coefficient of anionic chemicals is substantially lower than for the neutral form of the chemical, roughly a factor 10-100 for K_{OC} (Tülp et al. 2009). In case of weakly dissociating chemicals such as carboxylic acids, the sorption coefficient can often be estimated from the sorption coefficient of the non-ionic form and the fraction of the chemical that is present in the non-ionized form (see section on Relevant chemical properties).

Reliability and limitations of QSPR

Predictive models have limitations and it is important to know these limitations. There is not one single model that can predict a parameter for all chemicals. Each model will have a domain of applicability and it is important to apply a model only to a chemical within that domain. Therefore, guidance has to be defined on how to select a specific model. It is also important to realize that in many computer programs (such as fate modeling programs), estimates and predictions are implicitly incorporated in these progams.

Another aspect is the reliability of the prediction. The model itself can show a good fit (high r^2) for the training set (the chemicals used to develop the model), but the actual reliability should be tested with a separate set of chemicals (the validation set) and a

number of statistical procedures can be applied to test the accuracy and predictive power of the model. The OECD has developed a set of rules that should be applied in the validation of QSPR and QSAR models.

References

Abraham, M.H. (1993). Scales of solute hydrogen-bonding - their construction and application to physicochemical and biochemical processes. Chemical Society Reviews 22, 73-83.

Armitage, J.M., Arnot, J.A., Wania, F., Mackay, D. (2013). Development and evaluation of a mechanistic bioconcentration model for ionogenic organic chemicals in fish. Environmental Toxicology and Chemistry 32, 115-128.

Bruggeman, W.A., Opperhuizen, A., Wijbenga, A., Hutzinger, O. (1984). Bioaccumulation of super-lipophilic chemicals in fish. Toxicological and Environmental Chemistry 7, 173-189.

Droge, S.T.J., Goss, K.U. (2013). Development and evaluation of a new sorption model for organic cations in soil: Contributions from organic matter and clay minerals. Environmental Science and Technology 47, 14233-14241.

Endo, S., Escher, B.I., Goss, K.U. (2011). Capacities of membrane lipids to accumulate neutral organic chemicals. Environmental Science and Technology 45, 5912-5921.

Endo, S., Goss, K.U. (2014). Applications of polyparameter linear free energy relationships in environmental chemistry. Environmental Science and Technology 48, 12477-12491.

Eriksson, L., Hermens, J.L.M., Johansson, E., Verhaar, H.J.M., Wold, S. (1995). Multivariate analysis of aquatic toxicity data with pls. Aquatic Sciences 57:217-241.

Geisler, A., Endo, S., Goss, K.U. (2012). Partitioning of organic chemicals to storage lipids: Elucidating the dependence on fatty acid composition and temperature. Environmental Science and Technology 46, 9519-9524.

Goss, K.-U., Schwarzenbach, R.P. (2001). Linear free energy relationships used to evaluate equilibrium partitioning of organic compounds. Environmental Science and Technology 35, 1-9.

Goss, K.U., Schwarzenbach, R.P. (2003). Rules of thumb for assessing equilibrium partitioning of organic compounds: Successes and pitfalls. Journal of Chemical Education 80, 450-455.

Hansch, C., Streich, M., Geiger, F., Muir, R.M., Maloney, P.P., Fujita, T. (1963). Correlation of biological activity of plant growth regulators and chloromycetin derivatives with hammett constants and partition coefficients. Journal of the American Chemical Society 85, 2817-&.

Jonker, M.T.O., van der Heijden, S.A. (2007). Bioconcentration factor hydrophobicity cutoff: An artificial phenomenon reconstructed. Environmental Science and Technology 41, 7363-7369.

Katritzky, A.R., Slavov, S., Radzvilovits, M., Stoyanova-Slavova, I., Karelson, M. (2009). Computational chemistry approaches for understanding how structure determines properties. Zeitschrift Fur Naturforschung Section B-a Journal of Chemical Sciences 64:773-777.

Könemann, H., Van Leeuwen, K. (1980). Toxicokinetics in fish: Accumulation and elimination of six chlorobenzenes by guppies. Chemosphere 9, 3-19.

Kraaij, R., Mayer, P., Busser, F.J.M., Bolscher, M.V., Seinen, W., Tolls, J. (2003). Measured pore-water concentrations make equilibrium partitioning work - a data analysis. Environmental Science and Technology 37, 268-274.

Sabljic, A., Güsten, H., Verhaar, H.J.M., Hermens, J.L.M. (1995). Qsar modelling of soil sorption. Improvements and systematics of log k_{oc} vs. Log k_{ow} correlations. Chemosphere 31, 4489-4514.

Tülp, H.C., Fenner, K., Schwarzenbach, R.P., Goss, K.U. (2009). pH-dependent sorption of acidic organic chemicals to soil organic matter. Environmental Science and Technology 43, 9189-9195.

Veith, G.D., Defoe, D.L., Bergstedt, B.V. (1979). Measuring and estimating the bioconcentration factor of chemicals in fish. Journal of the Fisheries Research Board of Canada 36, 1040-1048.

3.4.3. Question 1

What is a QSPR and why is it useful?

3.4.3. Question 2

Which techniques are applied to derive a QSPR.

3.4.3. Question 3

Which chemical parameters are applied in a QSPR?

This page titled 2.3.3.4: Partitioning and Partitioning Constants is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **3.4: Partitioning and Partitioning Constants** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.3.5: Metal Speciation

Learning Objectives

You should be able to:

- describe the reactions involved in the speciation of metals in the aquatic and soil environments.
- explain the equilibrium approach to modelling metal speciation.
- identify which water and soil properties impact the fate of metals
- describe how processes such as competition and sorption impact metal bioavailability
- explain why the environmental fate of metals is dynamic

Keywords: Metal complexation, redox reactions, equilibrium reactions, water chemistry, soil properties.

Introduction

Metals occur in different physical and chemical forms in the environment, for example as the element (very rare in the environment), as components of minerals, as free cations dissolved in water (e.g. Cd^{2+}), or bound to inorganic or organic molecules in either the solid or dissolved phases (e.g. $HgCH_3^+$ or $AgCl^{2+}$) (Allen 1993). The distribution of a metal over these different forms is referred to as metal speciation. Physical processes may also affect the mobility and bioavailability of metals, for example the electrostatic attraction of metal cations to negatively charged mineral surfaces. These processes are in general not referred to as metal speciation in the strict sense but they are discussed here.

Metal speciation reactions

The speciation of metals is controlled by both the properties of the metals (see the section Metals and metalloids) and the properties of the environment in which they are present, such as the pH, redox potential and the presence and concentrations and properties of molecules that could form complexes with the metals. These complex forming molecules are often called ligands and these can vary from relatively simple anions in solution, such as sulphate or anions of soluble organic acids, to more complex macromolecules such as proteins and other biomolecules. The adsorption of metals by covalent bond formation to oxide and hydroxide surfaces of minerals, and oxygen- or nitrogen-containing functional groups of solid organic matter, is also referred to as complexation. Since these metal-binding functional groups are often either acidic or basic, the pH is an important environmental parameter controlling complexation reactions.

In natural systems the speciation of metals is of great complexity and determines their mobility in the environment and their bioavailability (i.e. how easily they are taken up by organisms). Metal speciation therefore plays a key role in determining the potential bioaccumulation and toxicity of metals and should therefore be considered when assessing their ecological risks. Metal bioavailability and transport are in particular strongly related to the distribution over solid and liquid phases of the environmental matrix.

The four main chemical reactions determining metal speciation, so the binding of metal ions to ligands and their presence in solid and liquid phases, are (Bourg, 1995):

- adsorption and desorption processes
- ion exchange and dissolution reactions
- precipitation and co-precipitation
- complexation to inorganic and organic ligands

The complexity of these reactions is illustrated in Figure 1.

Figure in preparation

Figure 1. *Metals (M) speciation in the environment is determined by a number of reactions, including complexation, precipitation and sorption. These reactions affect the partitioning of metals across solid and liquid phases, hence their mobility as well as their bioavailability.*

Adsorption, desorption and ion exchange processes take place with the reactive components present in soils, sediments and to lower extent in water. These include:

- clay minerals
- hydroxides (e.g. $FeOH_3$) and/or carbonates (e.g. $CaCO_3$)
- organic matter

Metal ions react with these reactive components in different ways. In soils and sediments, cationic metals bind reversibly to clay minerals via cation-exchange processes (see section on Soil). Metal ions also form complexes with so-called functional groups (mainly carboxylic and phenolic groups) present in organic matter (see section on Soil). In aquatic systems similar binding processes occur, in which dissolved organic matter (or carbon) (DOM or DOC) plays a major role. The "dissolved" fraction of organic matter is operationally defined as the fraction passing a 0.45 µm filter and is often referred to fulvic and humic acids.

As mentioned above and in the section on Shttps://maken.wikiwijs.nl/147644/Environmental_Toxicology__an_open_online_textbook#!page-5415168oil, negatively charged surfaces of the reactive mineral and organic components present in soil, sediment or water attract positively-charged atoms or molecules (cations, e.g. Cd^{2+}), and allow these cations to exchange with other positively charged ions. The competition between cations for binding sites is driven by the binding affinity of each metal species, as well as the concentration of each metal species. Cation-exchange capacity (CEC) is a property of the sorbent, and defined as the density of available negatively charged sites per mass of environmental matrix (soil, sediment). In fact, it is a measure of how many cations can be retained on solid surfaces. CEC usually is expressed in $cmol_c$/kg soil (see section on Soil). Increasing the pH (i.e. decreasing the concentration of H^+ ions) increases the variable charge of most sorbents (more types of protonated groups on sorbent surfaces release their H^+), especially for organic matter, and therefore also increases the cation exchange capacity. Protons (H^+) also compete with metal ions for the same binding sites. Conversely, at decreasing pH (increasing H^+ concentrations), most sorbents lower their CEC.

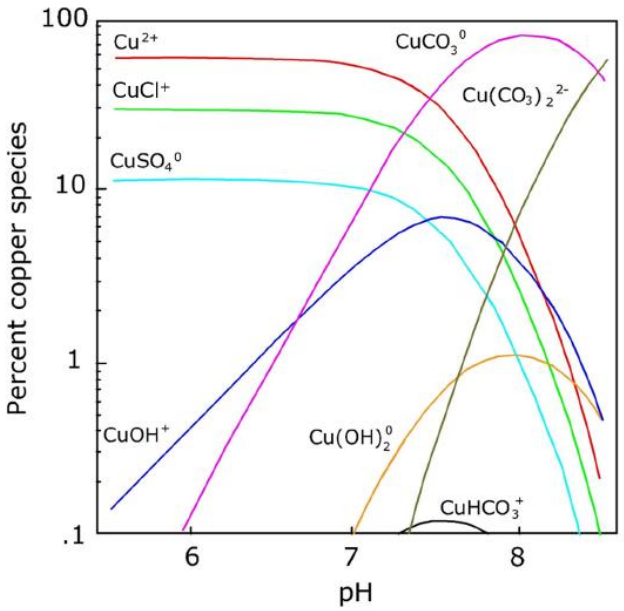
Modelling metal speciation

Metal speciation can be modelled if we have sufficient knowledge of the most important reactions involved and the environmental conditions that control these reactions. This knowledge is expressed in the form of equilibria expressing the most important complexation and/or redox reactions. For example, in the general case of a complexation reaction between metal M and ligand L described by the equilibrium:

aM^{m+} (aq) + bL^{n-} (aq) ↔ $M_aL_b^{q+}$ (aq) (where q = am-bn)

The relationship between the concentrations (or more accurately the activities) of the species is given by:

$$K_f = \frac{[M_a\ L_b^{q+}]}{[M^{m+}]^a\ +\ [L^{n-}]^b}$$ Figure 2. At pH 7.5, most Cu is bound to CO_3^{2-}, but at pH 6 the Cu is mainly present as the free Cu^{2+} ion and in



complexes with chloride (as $CuCl^+$) and sulphate (as $CuSO_4$).
Figure 2. Copper speciation in salt water without DOC. Redrawn from Blust et al. (1991) by Wilma Ijzerman.

If redox reactions are involved in speciation, we can use the Nernst equation to describe the equilibrium between reduced and oxidised states of the metal:

$$E_h = E_h^0 - \frac{2.303RT}{nF} log \frac{(Red))}{(Ox)}$$

Where E_h is the redox potential, E_h^0 the standard potential of the redox pair (relative to the hydrogen electrode), R the molar gas constant, T the temperature, n the number of transferred electrons, F the Faraday constant and {Red/Ox} the activity (or concentration) ratio of the reduced and oxidized species. Since many redox reactions involve the transfer of H^+, the value of {Red/Ox} for these equilibria will depend on the pH. Note that the redox potential is often expressed as p*e* which is defined as the negative logarithm of the electron activity (p*e* = - log $\{e^-\}$).

Using these comparatively simple equations for all the relevant reactions involved it is possible to construct models to describe metal speciation as a function of ligand concentrations, pH and redox potential. As an example, Table 1 presents the relevant equilibria for the speciation of iron in water.

Table 1. Equilibrium reactions relevant for Fe in water (adapted from Essington, 2003)

Boundary	**Equilibrium reaction**	
(1)	$Fe^{3+} + e^-$ D Fe^{2+}	$pE^\Theta = 13.05$
(2)	$Fe(OH)_3(s) + 3H^+$ D $Fe^{3+} + 3H_2O$	$K_{sp} = 9.1 \times 10^3\ L^2\ mol^{-2}$
(3)	$Fe(OH)_2(s) + 2H^+$ D $Fe^{2+} + 2H_2O$	$K^*_{sp} = 8.0 \times 10^{12}\ L\ mol^{-1}$
(4)	$Fe(OH)_3(s) + H^+ + e^-$ D $Fe(OH)_2(s) + H_2O$	
(5)	$Fe(OH)_3(s) + 3H^+ + e^-$ D $Fe^{2+} + 3H_2O$	

Using these we can derive equations defining the conditions of pH and p*e* at which the activity or concentration ratio is one for each equilibrium. These are shown as the continuous boundary lines in Fig. 3. In this Pourbaix or p*e*-pH (or pE-pH) diagram, the fields separated by the boundary lines are labelled with the dominant species present under the conditions that define the fields. (NB. The dotted lines define the conditions of p*e* and pH under which water is stable.)

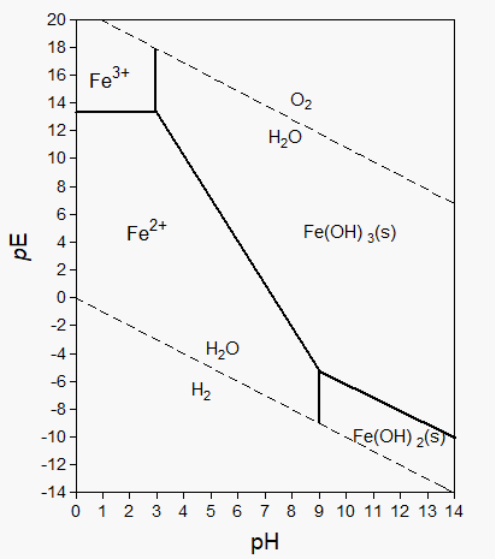


Figure 3. A pe-pH diagram for Fe in water showing the dominant species present under different conditions assuming a maximum soluble Fe(II) or (III) concentration of 10^{-5} mol L^{-1} (adapted from Essington, 2003).

Environmental effects on speciation

In the environment there is, however, in general no equilibrium. This means that the speciation and hence also fate of metals is highly dynamic. Large scale alterations occur when land use changes, e.g. when agricultural land is abandoned and becomes nature. Whereas agricultural soil often is 'limed' (addition of $CaCO_3$) to maintain near-neutral pH and crop is removed by harvesting, in natural ecosystems all produced organic matter remains in the system. Therefore natural soils show an increase in soil organic matter content, while due to microbial decomposition processes soil pH tends to decrease. As a result, DOC concentration in the soil porewater will increase, while metal mobility also is increased by the decreasing soil pH (Cu^{2+} is more mobile than $CuCO_3$). This may cause historical metal pollution to suddenly become available (the "chemical time bomb" effect). Large scale reconstruction of rivers or deep soil digging for land planning and development may also affect environmental conditions in such a way that metal speciation may change. An example of this is the change in arsenic speciation in groundwater due to the drilling of wells in countries like Bangladesh; the introduction of oxygen and organic matter into the deeper groundwater caused a change of arsenic speciation, enhancing its solubility in water and therefore increasing human exposure (see section on Metals and metalloids).

Dynamic conditions do not only occur on large spatial and temporal scales, nature is also dynamic on smaller scales. Abiotic factors such as rain and flooding events, weather conditions, and redox status may alter metal speciation. In addition, biotic factors may affect metal speciation. An example of the latter is the bioturbation by sediment-dwelling organisms that re-suspend particles into water, or earthworms that by their digging activities aerate the soil and excrete mucus that may stimulate microbial activity (see Figure 4A). These activities of soil and sediment organisms alter the environmental conditions and hence affect metal speciation (see Figure 4B). The production of acidic root exudates by plants may also have similar effects on metal speciation. Another process that alters metal speciation is the uptake of metals. Since the ionic metal form seems most prone to root uptake, or active intake over cell membranes, this process may affect metal partitioning over different species.

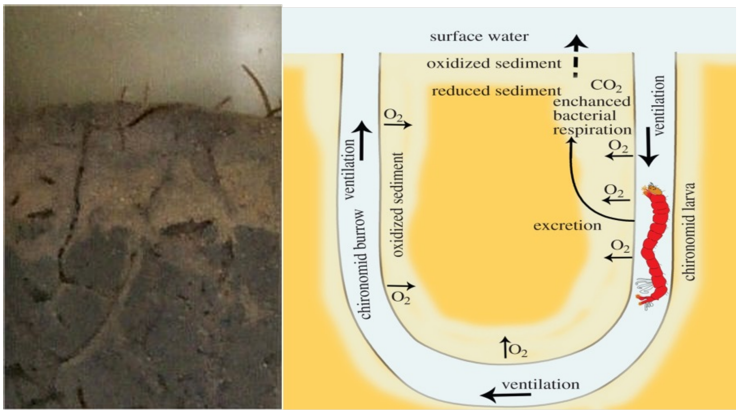


Figure 4. Illustration of small scale processes that alter metal speciation. (A) different bioturbation activities by various organisms, (B) the burrowing activities of a chironomid (midge) larvae can alter the environmental conditions and with that affect metal speciation in the sediment. Source: Martina Vijver.

References

Allen, H.E. (1993). The significance of trace metal speciation for water, sediment and soil quality criteria and standards. Science of the Total Environment 134, 23-45.

Andrews, J.E., Brimblecombe, P., Jickells, T.D., Liss P.S., Reid, B. (2004) An Introduction to Environmental Chemistry, 2nd Edition, Blackwell, ISBN 0-632-05905-2 (chapter 6).

Bourg, A.C.M. (1995) Speciation of heavy metals in soils and groundwater and implications for their natural and provoked mobility. In: Salomons, W., Förstner, U., Mader, P. (Eds.). Heavy Metals. Springer, Berlin. p. 19-31.

Blust, R., Fontaine, A., Decleir, W. (1991) Effect of hydrogen ions and inorganic complexing on the uptake of copper by the brine shrimp *Artemia franciscana*. Marine Ecology Progress Series 76, 273-282.

Essington, M.E. (2003) Soil and Water Chemistry, CRC Press, ISBN 0-8493-1258-2 (chapters 5, 7 and 9).

Sposito, G. (2008) The Chemistry of Soils, 2nd Edition, Oxford University press, ISBN 978-0-19-531369-7 (chapter 4).

Sparks, D.L. (2003) Environmental Soil Chemistry, 2nd Edition, Academic Press, ISBN 0-12-656446-9 (chapters 5, 6 and 8).

3.5. Question 1

What are the most important reactions involved in the speciation of metals in the aquatic and soil environments?

3.5. Question 2

What are the most important environmental parameters controlling these reactions?

3.5. Question 3

In the equilibrium approach to modelling metal speciation, dominance or p*e*-pH diagrams are used as a visual representation of speciation. What do the lines and fields in such diagrams represent?

Contributors

- **Author:** Martina Vijver, John Parsons
- **Reviewers:** Kees van Gestel, Ronny Blust, Steven Droge

This page titled 2.3.3.5: Metal Speciation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **3.5: Metal Speciation** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

SECTION OVERVIEW

2.3.4: Toxicology

2.3.4.1: Toxicokinetics

2.3.4.2: Toxicodynamics and Molecular Interactions

2.3.4.3: Toxicity Testing

2.3.4.4: Increasing ecological realism in toxicity testing

2.3.4.5: Availability and bioavailability

2.3.4.6: Degradation

2.3.4.7: Modelling exposure

2.3.4.8: Section 4.1. Toxicokinetics

2.3.4.9: Section 4.4. Increasing ecological realism in toxicity testing

2.3.4.10: Section 4.2. Toxicodynamics and Molecular Interactions

2.3.4.11: Section 4.3. Toxicity testing

This page titled 2.3.4: Toxicology is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

2.3.4.1: Toxicokinetics

4.1. Toxicokinetics

Author: Nico van Straalen

Reviewer: Kees van Gestel

Learning objectives:

You should be able to

- Describe the difference between toxicokinetics and toxicodynamics
- Explain the use of different descriptors for the uptake of chemicals by organisms under steady state and dynamic conditions

Keywords: toxicodynamics, toxicokinetics, bioaccumulation, toxicokinetic rate constants, critical body residue, detoxification

Toxicology usually distinguishes between toxicokinetics and toxicodynamics. Toxicokinetics involves all processes related to uptake, internal transport and the accumulation inside an organism, while toxicodynamics deals with the interaction of a compound with a receptor, induction of defence mechanisms, damage repair and toxic effects. Of course the two sets of processes may interact, for instance defence may feed-back to uptake and damage may change the internal transport. However, often toxicokinetics analysis just focuses on tracking of the chemical itself and ignores possible toxic effects. This holds up to a critical threshold, the critical body concentration, above which effects become obvious and the normal toxicokinetics analysis is no longer valid. The assumption that toxicokinetic rate parameters are independent of the internal concentration is due the limited amount of information that can be obtained from animals in the environment. However, in so called *physiology-based pharmacokinetic and pharmacodynamic models* (PBPK models) kinetics and dynamics are analyzed as an integrated whole. The use of such models is however mostly limited to mammals and humans.

It must be emphasized that toxicokinetics considers fluxes and rates, i.e. mg of a substance moving per time unit from one compartment to another. Fluxes may lead to a dynamic equilibrium, i.e. an equilibrium that is due to inflow being equal to outflow; when only the equilibrium conditions are considered, this is called partitioning.

In this Chapter 4.1 we will explore the various approaches in toxicokinetics, including the fluxes of toxicants through individual organisms and through trophic levels as well as the biological processes that determine such fluxes. We start by comparing the concentrations of toxicants between organisms and their environment (section 4.1.1), and between organisms of different trophic levels (section 4.1.6). This leads to the famous concept of bioaccumulation, one of the properties of a substance often leading to environmental problems. While in the past dilution was sometimes seen as a solution to pollution, this is not correct for bioaccumulating substances, since they may turn up elsewhere in the next level food-chain and reach an even higher concentration. The bioaccumulation factor is one of the best-investigated properties characterizing the environmental behaviour of a substance . It may be predicted from properties of the substance such as the octanol-water partitioning coefficient.

In section 4.1.2 we discuss the classical theory of uptake-elimination kinetics using the one-compartment linear model. This theory is a crucial part of toxicological analysis. One of the first things you want to know about a substance is how quickly it enters an organism and how quickly it is removed. Since toxicity is basically a time-dependent process, the turnover rate of the internal concentration and the build-up of a residue depend upon the exposure time. An understanding of toxicokinetics is therefore critical to any interpretation of a toxicity experiment. Rate parameters may partly be predicted from substance properties, but properties of the organism play a much greater role here. One of these is simply the body mass; prediction of elimination rate constants from body mass is done by allometric scaling relationships, explored in section 4.1.5.

In two sections, 4.1.3 and 4.1.4, we present the biological processes that underlie the turnover of toxicants in an organism. These are very different for metals than for organic substances, hence, two separate sections are devoted to this topic, one on tissue accumulation of metals and one on defence mechanisms for organic xenobiotics.

Finally, if we understand all toxicokinetic processes we will also be able to understand whether the concentration inside a target organ will stay below or just passes the threshold that can be tolerated. The critical body concentration, explored in section 4.1.7 is an important concept linking toxicokinetics to toxicity.

4.1. Question 1

Define and indicate the differences between

1. Toxicokinetics
2. Toxicodynamics
3. Partitioning
4. Accumulation
5. Bioaccumulation

4.1.1. Bioaccumulation

Author: Joop Hermens

Reviewers: Kees van Gestel and Philipp Mayer

Learning objectives:

You should be able to

- define and explain different bioaccumulation parameters.
- Mention different biological factors that may affect bioaccumulation.

Key words: Bioaccumulation, lipid content

Introduction: terminology for bioaccumulation

The term bioaccumulation describes the transfer and accumulation of a chemical from the environment into an organism." For a chemical like hexachlorobenzene, the concentration in fish is more than 10,000 times higher than in water, which is a clear illustration of "bioaccumulation". A chemical like hexachlorobenzene is hydrophobic, so has a very low aqueous solubility. It therefore prefers to escape the aqueous phase to enter (or partition into) a more lipophilic phase such as the lipid phase in biota.

Uptake may take place from different sources. Fish mostly take up chemicals from the aqueous phase, organisms living at the sediment-water interphase are exposed via the overlying water and sediment particles, organisms living in soil or sediment via pore water and by ingesting soil or sediment, while predators will be exposed via their food. In many cases, uptake is related to more than one source. The different uptake routes are also reflected in the parameters and terminology used in bioaccumulation studies. The different parameters include the bioconcentration factor (BCF), bioaccumulation factor (BAF), biomagnification factor (BMF) and biota-to-sediment or biota-to-soil accumulation factor (BSAF). Figure 1 summarizes the definition of these parameters. Bioconcentration refers to uptake from the aqueous phase, bioaccumulation to uptake via both the aqueous phase and the ingestion of sediment or soil particles, while biomagnification expresses the accumulation of contaminants from food.

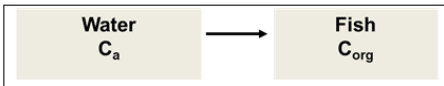
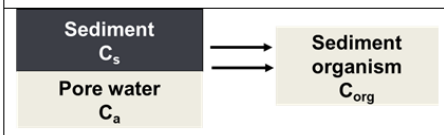
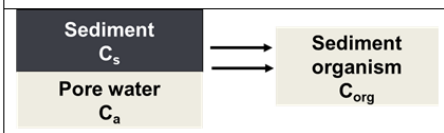
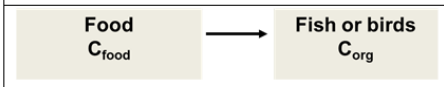
	Uptake from water only	Bioconcentration factor	BCF (L/kg)	C_{org} / C_{aq}
	Uptake from water and sediment or soil	Bioaccumulation factor	BAF (L/kg)	C_{org} / C_{aq}
	Uptake from water and sediment or soil	Biota-to-sediment or Biota-to-soil accumulation factor	BSAF (kg soil / kg org.)	C_{org} / C_{sed} or C_{org} / C_{soil}
	Uptake from food	Biomagnification factor	BMF (kg food/ kg org.)	C_{org} / C_{food}

Figure 1. Parameters used to describe the bioaccumulation of chemicals.

C_{aq} concentration in water (aqueous phase)

C_{org} concentration in organism

C_f concentration in food

C_s concentration in sediment or soil

Please note that the bioaccumulation factor (BAF) is defined in a similar way as the bioconcentration Factor (BCF), but that uptake can be both from the aqueous phase and the sediment or soil and that the exposure concentration usually is expressed per kg dry sediment or soil. Other definitions of the BAF are possible, but we have followed the one from Mackay et al. (2013). "The bioaccumulation factor (BAF) is defined here in a similar fashion as the BCF; in other words, BAF is CF/CW at steady state, except that in this case the fish is exposed to both water and food; thus, an additional input of chemical from dietary assimilation takes place".

All bioaccumulation factors are steady state constants: the concentration in the organism constant and the organisms is in equilibrium with its surrounding phase. It will take time before such an steady state is reached. Steady state is reached when uptake rate (for example from an aqueous phase) equals the elimination rate. Models that include the factor time in describing the uptake are called kinetic models; see section on Bioaccumulation kinetics.

Effect of biological properties on accumulation

Uptake of chemicals is determined by properties of both the organism and the chemical. For xenobiotic lipophilic chemicals in water, organism-specific factors usually play a minor role and concentrations in organisms can pretty well be predicted from chemical properties (see section on Structure-property relationships). For metals, on the contrary, uptake is to a large extent determined by properties of the organism, and a direct consequence of its mineral requirements. A chemical with low bioavailability (low uptake compared to concentration in exposure medium) may nevertheless accumulate to high levels when the organism is not capable of excreting or metabolising the chemical.

Factors related to the organism are:

- Fat content. Because lipophilic chemicals mainly accumulate in the fat of organisms, it is reasonable to assume that lipid rich organisms will have higher concentrations of lipophilic chemicals. See Figure 2 for an example of bioconcentration factors of 1,2,4-trichlorobenzene in a number of organisms with varying lipid content. This was one of the explanations for the high PCB levels in eel (high lipid content) in Dutch rivers and seals from the Wadden Sea. Nevertheless, large differences are still found between species when concentrations are expressed on a lipid basis. This may e.g. be explained by the fact that lipids are not all identical: PCBs seem to dissolve better in lipids from anchovies than in lipids from algae (see data in Table 1). Also more recent research has shown that not all lipids are the same and that differences in bioaccumulation between species may be due to differences in lipid composition (Van der Heijden and Jonker, 2011). Related to this is the development of BCF models that are based on multiple compartments and that make a separation between storage and membrane lipids and also include a protein fraction as additional sink (Armitage et al., 2013). In this model, the overall distribution coefficient D_{BW} (or BCF) is estimated via equation 1. The equation is using distribution coefficients because the model also "accounts for the presence of neutral and charged chemical species" (Armitage et al., 2013).

$$BCF \quad or \quad D_{BW} = f_{SL} D_{SL-W} + f_{ML} D_{ML-W} + f_{NLOM} D_{NLOM-W} + f_W$$

D_{BW}	overall organism-water distribution coefficient (or surrogate BCF) at a given pH
D_{SL-W}	storage lipid-water distribution ratio
D_{ML-W}	membrane lipid-water distribution ratio
D_{NLOM-W}	sorption coefficient to NLOM (non-lipid organic matter, for example proteins)
f_{SL}	fraction of storage lipids
f_{ML}	fraction of membrane lipids
f_{NLOM}	fraction of non-lipid organic matter (e.g. proteins, carbohydrates)
f_W	fraction of water

- Sex: Chemicals (such as DDT, PCB) accumulated in milk fat may be transferred to juveniles upon lactation. This was found in marine mammals. In this way, females have an additional excretion mechanism. A typical example is shown in Figure 3, taken from a study of Abarnou et al. (1986) on the levels of organochlorinated compounds in the Antarctic dolphin *Cephalorhyncus commersonii*. In males, concentrations increase with increasing age, but concentrations in mature females decrease with increasing age.
- Weight: (body mass) of the organism relative to the surface area across which exchange with the water phase takes place. Smaller organisms have a larger surface-to-volume ratio and the exchange with the surrounding aqueous phase is faster. Therefore, although lipid normalized concentrations will be the same at equilibrium, this equilibrium is reached earlier in smaller than in larger organisms.
- Difference in uptake route. The relative importance of uptake through the skin (in fish e.g. the gills) and (oral) uptake through the digestive system. It is generally accepted that for most free-living organism's direct uptake from the water dominates over uptake after digestion in the digestive tract.
- Metabolic activity. Even at the same weight or the same age, the balance between uptake and excretion may change due to an increased metabolic activity, e.g. in times of fast growth or high reproductive activity.

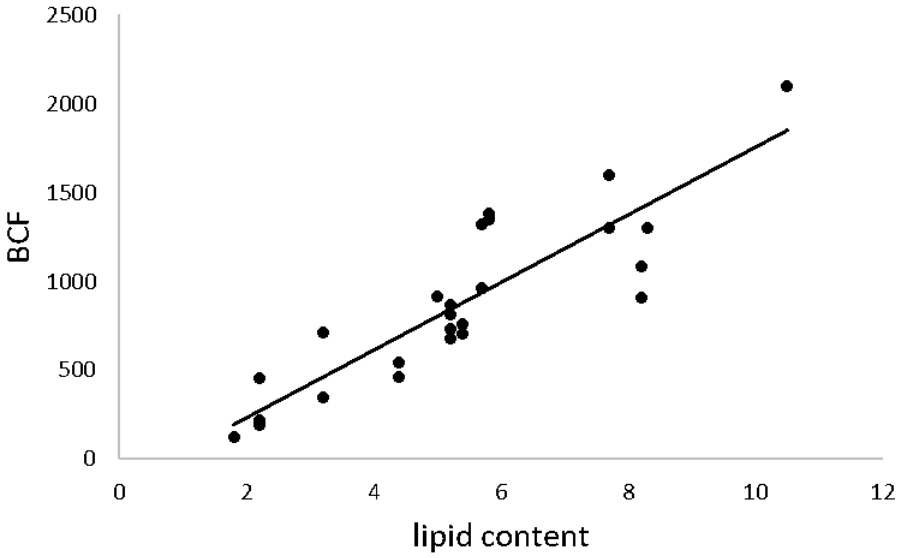


Figure 2. The influence of lipid content on the bioconcentration of 1,2,4-trichlorobenzene in different fish species (reproduced using data from Geyer et al., 1985).

Table 1. *Mean PCB concentrations in algae (Dunaliella spec.), rotifers (Brachionus plicatilis) and anchovies larvae (Angraulis mordax), expressed on a dry-weight basis and on a lipid basis. From Moriarty (1983).*

Organism	Lipid content (%)	PCB-concentration based on dry weight (µg g^{-1})	PCB-concentration based on lipid weight (µg g^{-1})	BCF based on concentration in the lipid phase
algae	6.4	0.25	3.91	0.48 x 10^6
rotifer	15.0	0.42	2.80	0.34 x 10^6
fish (anchovies) larvae	7.5	2.06	27.46	13.70 x 10^6

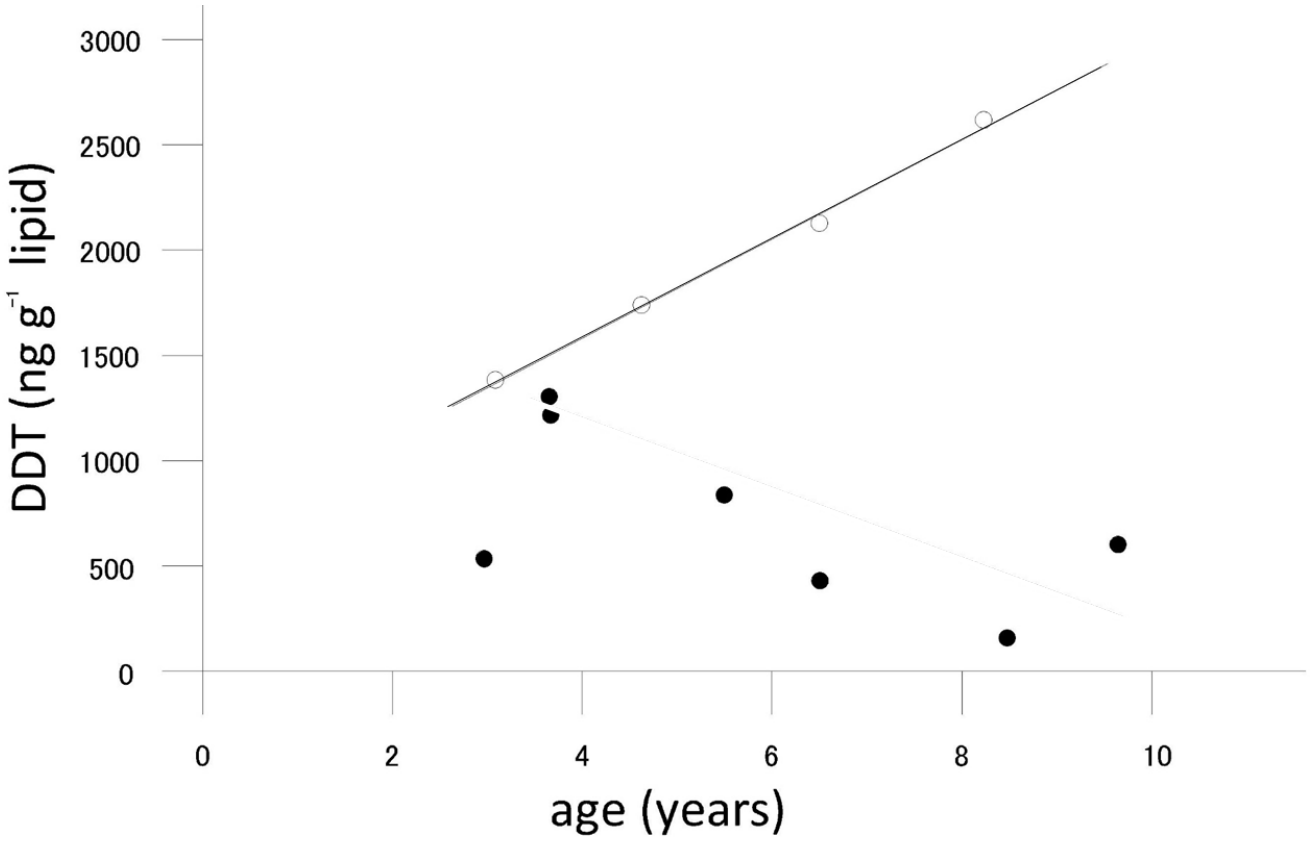


Figure 3. Concentrations of DDT in dolphins of different age and the difference between male and female dolphins Redrawn from Abernou et al. (1986) by Wilma IJzerman.

Cited references

Abarnou, A., Robineau, D., Michel, P. (1986). Organochlorine contamination of commersons dolphin from the Kerguelen islands. Oceanologica Acta 9, 19-29.

Armitage, J.M., Arnot, J.A., Wania, F., Mackay, D. (2013). Development and evaluation of a mechanistic bioconcentration model for ionogenic organic chemicals in fish. Environmental Toxicology and Chemistry 32, 115-128.

Geyer, H., Scheunert, I., Korte, F. (1985). Relationship between the lipid-content of fish and their bioconcentration potential of 1,2,4-trichlorobenzene. Chemosphere 14, 545-555.

Mackay, D., Arnot, J.A., Gobas, F., Powell, D.E. (2013). Mathematical relationships between metrics of chemical bioaccumulation in fish. Environmental Toxicology and Chemistry 32, 1459-1466.

Moriarty, F. (1983). *Ecotoxicology: The Study of Pollutants in Ecosystems.* Publisher: Academic Press, London.

Van der Heijden, S.A., Jonker, M.T.O. (2011). Intra- and interspecies variation in bioconcentration potential of polychlorinated biphenyls: are all lipids equal? Environmental Science and Technology 45, 10408-10414.

Suggested reading

Mackay, D., Fraser, A. (2000). Bioaccumulation of persistent organic chemicals: Mechanisms and models. Environmental Pollution 110, 375-391.

Van Leeuwen, C.J., Vermeire, T.G. (Eds.) (2007). Risk Assessment of Chemicals: An Introduction. Springer, Dordrecht, The Netherlands. Chapter 3.

4.1.1. Question 1

What is the difference between BCF and BSAF?

4.1.1. Question 2

What are the main uptake routes for a sediment organism

4.1.1. Question 3

Which biological factors may influence the bioaccumulation?

4.1.2. Toxicokinetics

Author: Joop Hermens, Nico van Straalen

Reviewers: Kees van Gestel, Philipp Mayer

Learning objectives:

You should be able to

- mention the underlying assumptions of the kinetic models for bioaccumulation
- understand the basic equations of a one compartment kinetic bioaccumulation model
- explain the differences between one- and two-compartment models
- mention which factors affect the rate constants in a compartment model

Key words: Bioaccumulation, toxicokinetics, compartment models

In the section "Bioaccumulation", the process of bioaccumulation is presented as a steady state process. Differences in the bioaccumulation between chemicals are expressed via, for example, the bioconcentration factor BCF. The BCF represents the ratio of the chemical concentration in, for instance, a fish versus the aqueous concentration at a situation where the concentrations in water and fish do not change in time.

$$BCF = \frac{Corg}{Caq} \quad (1)$$

where:

C_{aq} concentration in water (aqueous phase) (mg/L)

C_{org} concentration in organism (mg/kg)

The unit of BCF is L/kg.

Kinetic models

Steady state can be established in a simple laboratory set-up where fish are exposed to a chemical at a constant concentration in the aqueous phase. From the start of the exposure (time 0, or t=0), it will take time for the chemical concentration in the fish to reach steady state and in some cases, this will not be established within the exposure period. In the environment, exposure concentrations may fluctuate and, in such scenarios, constant concentrations in the organism will often not be established. Steady state is reached when the uptake rate (for example from an aqueous phase) equals the elimination rate. Models that include the factor time in describing the uptake of chemicals in organisms are called kinetic models.

Toxicokinetic models for the uptake of chemicals into fish are based on a number of processes for uptake and elimination. An overview of these processes is presented in Figure 1. In the case of fish, the major process of uptake is by diffusion from the surrounding water compartment via the gill to the blood. Elimination can be via different processes: diffusion via the gill from blood to the surrounding water compartment, via transfer to offspring or eggs by reproduction, by growth (dilution) and by internal degradation of the chemical (biotransformation).

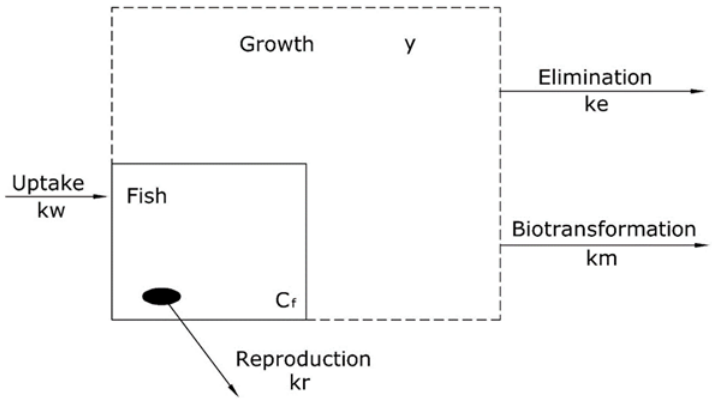


Figure 1. Uptake and elimination processes in fish and the rate constants (k) for each process. Reproduced from Van Leeuwen and Vermeire (2007) by Wilma IJzerman.

Kinetic models to describe uptake of chemicals into organisms are relatively simple with the following assumptions:

First order kinetics:

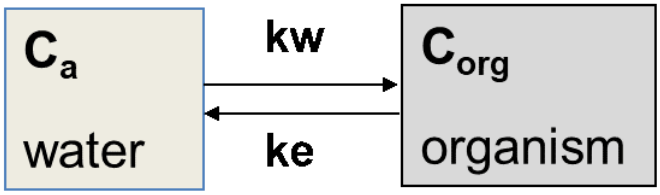
Rates of exchange are proportional to the concentration. The change in concentration with time (dC / dt) is related to the concentration and a rate constant (k):

$$\frac{dC}{dt} = kC \quad (2)$$

One compartment:

It is often assumed that an organism consists of only one single compartment and that the chemical is homogeneously distributed within the organism. For "simple" small organisms this assumption is intuitively valid, but for large fish this assumption looks unrealistic. But still, this simple model seems to work well also for fish. To describe the internal distribution of a chemical within fish, more sophisticated kinetic models are needed, similar to the ones applied in mammalian studies. These more complex models are the "physiologically based toxicokinetic" (PBTK) models (Clewell, 1995; Nichols et al., 2004)

Equations for the kinetics of accumulation process



The accumulation process can be described as the sum of rates for uptake and elimination.

$$\frac{dC_{org}}{dt} = k_w C_{aq} - k_e C_{org} \quad (3)$$

Integration of this differential equation leads to equation 4.

$$C_{org}(t) = \frac{k_w}{k_e} C_{aq}(1 - e^{k_e t}) \quad (4)$$

C_{org} concentration in organisms (mg/kg)

C_{aq} concentration in aqueous phase (mg/L)

k_w uptake rate constant (L/kg·day)

k_e elimination rate constant (1/day)

t time (day)

(dimensions used are: amount of chemical: mg; volume of water: L; weight of organism: kg; time: day); see box.

Box: The units of toxicokinetic rate constants

The differential equation underlying toxicokinetic analysis is basically a mass balance equation, specifying conservation of mass. A mass balance implies that the amount of chemical is expressed in absolute units such as mg. If Q is the amount in the animal and F the amount in the environmental compartment the mass balance reads:

$$\frac{dQ}{dt} = k_1' F(t) - k_2 Q(t)$$



where is the uptake rate constant and k_2 the elimination rate constant, both with dimension time^{-1}. However, it is often more practical to work with the concentration in the animal (e.g. expressed in mg/kg). This can be achieved by dividing the left and right sides of the equation by w, the body-weight of the animal and defining $C_{int} = Q/w$. In addition, we define the external concentration as $C_{env} = F/V$, where V is the volume (L or kg) of the environmental compartment. This leads to the following formulation of the differential equation:

$$\frac{dC_{int}}{dt} = k_1' \frac{V}{w} C_{env} - k_2 C_{int}$$

Beware that C_{env} is measured in other units (mg per kg of soil, or mg per litre of water) than C_{int} (mg per kg of animal tissue). To get rid of the awkward factor V/w it is convenient to define a new rate constant, k_1:

$$k_1 = \frac{V}{w} k_1'$$

https://bio.libretexts.org/@go/page/108678

This is the uptake rate constant usually reported in scientific papers. Note that it has other units than



: it is expressed as kg of soil per kg of animal tissue per time unit (kg kg^{-1} h^{-1}), and in the case of water exposure as L kg^{-1} h^{-1}. The dimension of k_2 remains the same whether mass or concentrations are used ($time^{-1}$). We also learn from this analysis that when dealing with concentrations, the body-weight of the animal must remain constant.

References

Moriarty, F. (1984) Persistent contaminants, compartmental models and concentration along food-chains. *Ecological Bulletins* **36**: 35-45.

Skip, B., A.J. Bednarska, & R. Laskowski (2014) Toxicokinetics of metals in terrestrial invertebrates: making things straight with the one-compartment principle. *PLoS ONE* **9**(9): e108740.

Equation 4 describes the whole process with the corresponding graphical representation of the uptake graph (Figure 2).

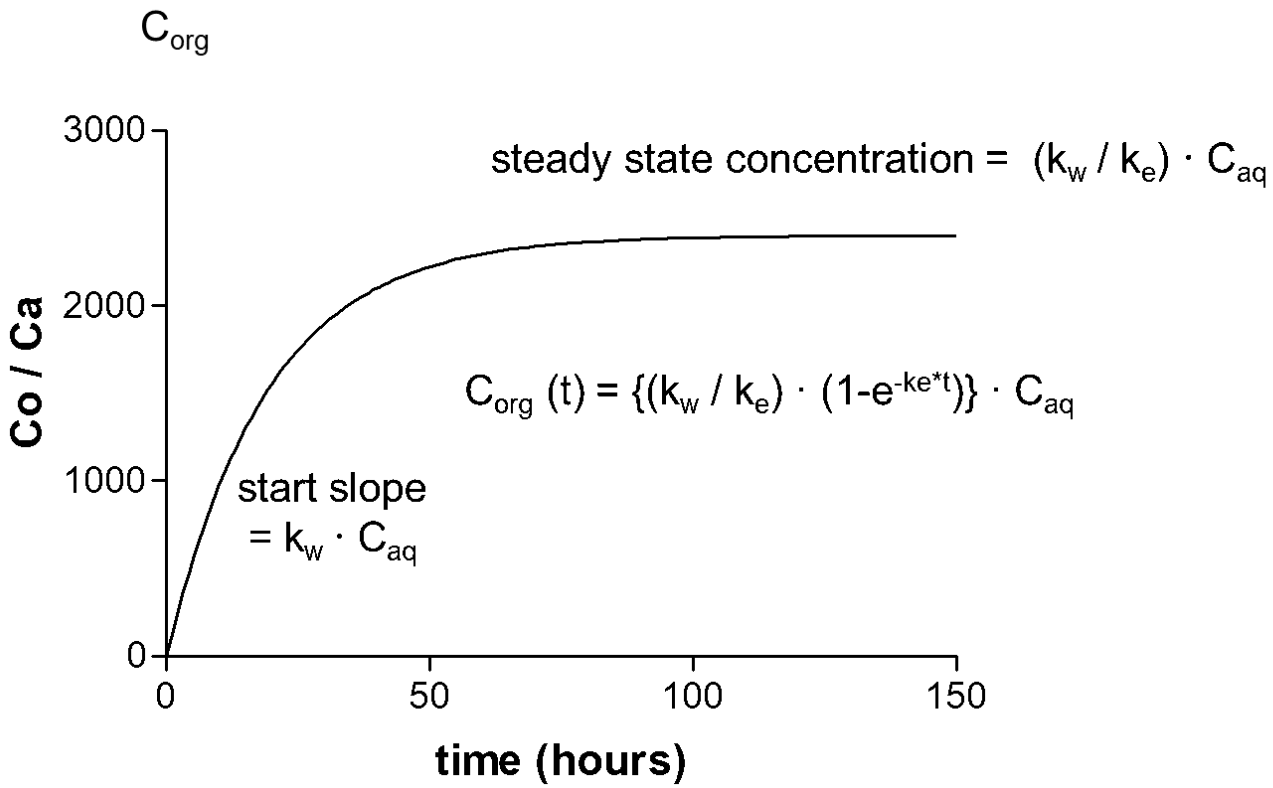


Figure 2. The basic equation for uptake of a chemical from the aqueous phase to a fish: one compartment model with first order kinetics.

The concentration in the organism is the result of the net process of uptake and elimination. At the initial phase of the accumulation process, elimination is negligible and the ratio of the concentration in the organism is given by:

$$\frac{dC_{org}}{dt} = k_w C_{aq} \quad (5)$$

$$C_{org}(t) = k_w C_{aq} t \quad (6)$$

Steady state

After longer exposure time, elimination becomes more substantial and the uptake curve starts to level off. At some point, the uptake rate equals the elimination rate and the ratio



C_{org}/C_{aq} becomes constant. This is the steady state situation. The constant C_{org}/C_{aq} at steady state is called the bioconcentration factor BCF. Mathematically, the BCF can also be calculated from



k_w/k_e. This follows directly from equation 4: after long exposure time (t), $e^{k_e t}$ becomes 0 leading to

$$C_{org}(t) = \frac{k_w}{k_e} C_{aq}$$

$$\frac{C_{org}}{C_{aq}}(at \quad steady \quad state) = BCF = \frac{k_w}{k_e} \quad (7)$$

Elimination

Elimination is often measured following an uptake experiment. After the organism has reached a certain concentration, fish are transferred to a clean environment and concentration in the organism will decrease in time. Because this is also a first order kinetic process, the elimination rate will depend on the concentration in the organism (C_{org}) and the elimination rate constant (k_e) (see equation 8). Concentration will decrease exponentially in time (equation 9) as shown in Figure 3A. Concentrations are often transformed to the natural logarithmic values (ln C_{org}) because this results in a linear relationship with slope -k_e.(equation 10 and figure 3B).

$$\frac{dC_{org}}{dt} = -k_e C_{org} \quad (8)$$

$$C_{org}(t) = C_{org}(t=0)e^{-k_e t} \quad (9)$$

$$lnC_{org}(t) = lnC_{org}(t=0) - k_e t \quad (10)$$

where C_{org}(t=0) is the concentration in the organism when the elimination phase starts.

The half-life ($T_{1/2}$ or DT_{50}) is the time needed to eliminate half the amount of chemical from the compartment. The relationship between k_e and $T_{1/2}$ is: $T_{1/2}$ = (In 2) / k_e. The half-life increases when k_e decreases.

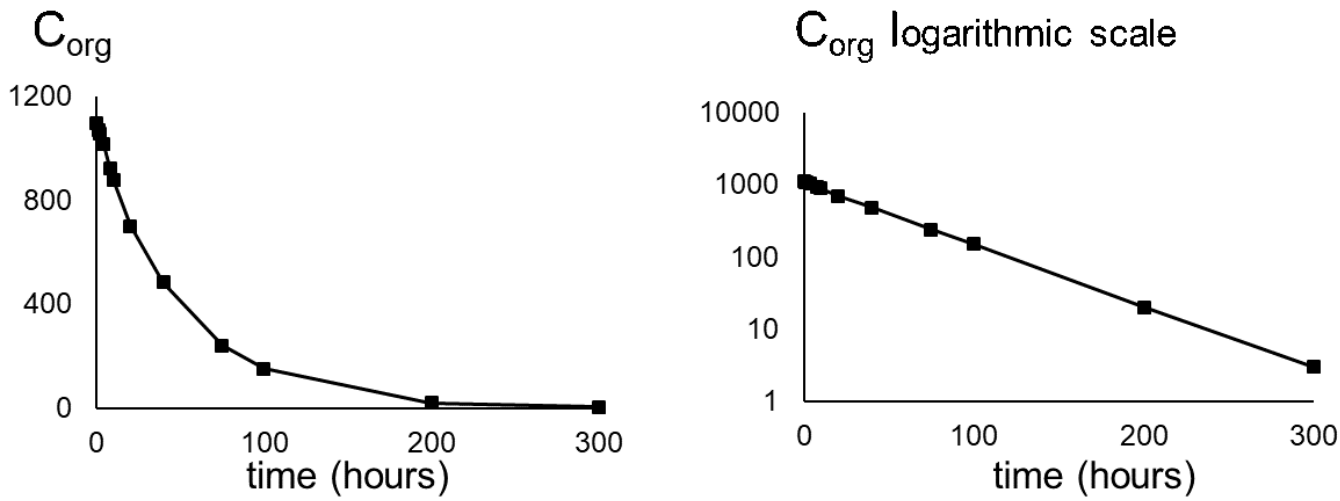


Figure 3. Elimination of a chemical from fish to water: one compartment model. Left: concentrations given on a normal scale. Right: concentration expressed as the natural logarithm to enable linear regression against time to yield the rate constant as the slope.

Multicompartment models

Very often, organisms cannot be considered as one compartment, but as two or even more compartments (Figure 4A). Deviations from the one-compartment system usually are seen when elimination does not follow an exponential pattern as expected: no linear relationship is obtained after logarithmic transformation. Figure 4B shows the typical trend of the elimination in a two-compartment system. The decrease in concentration (on a logarithmic scale) shows two phases: phase I with a relatively fast decrease and phase II with a relatively slow decrease. According to the linear compartment theory, elimination may be described as the sum of two (or more) exponential terms, like:

$$C_{org}(t) = C_{org}(t=0)F(I)e^{-k_e(I)*t} + F(II)e^{-k_e(II)*t} \quad (11)$$

where

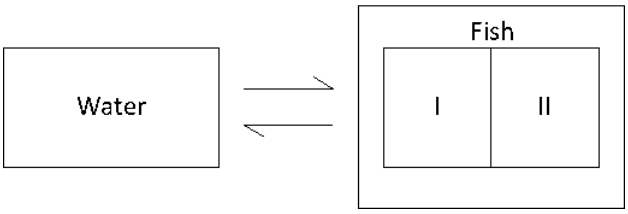
$k_e(I)$ and $k_e(II)$ represent elimination rate constants for compartment I and II,

F(I) and F(II) are the size of the compartments (as fractions)

Typical examples of two compartment systems are:

- Blood (I) and liver (II)
- Liver tissue (I) and fat tissue (II)

Elimination from fat tissue is often slower than from, for example, the liver. The liver is a well perfused organ while the exchange between lipid tissue and blood is much less. That explains the faster elimination from the liver.



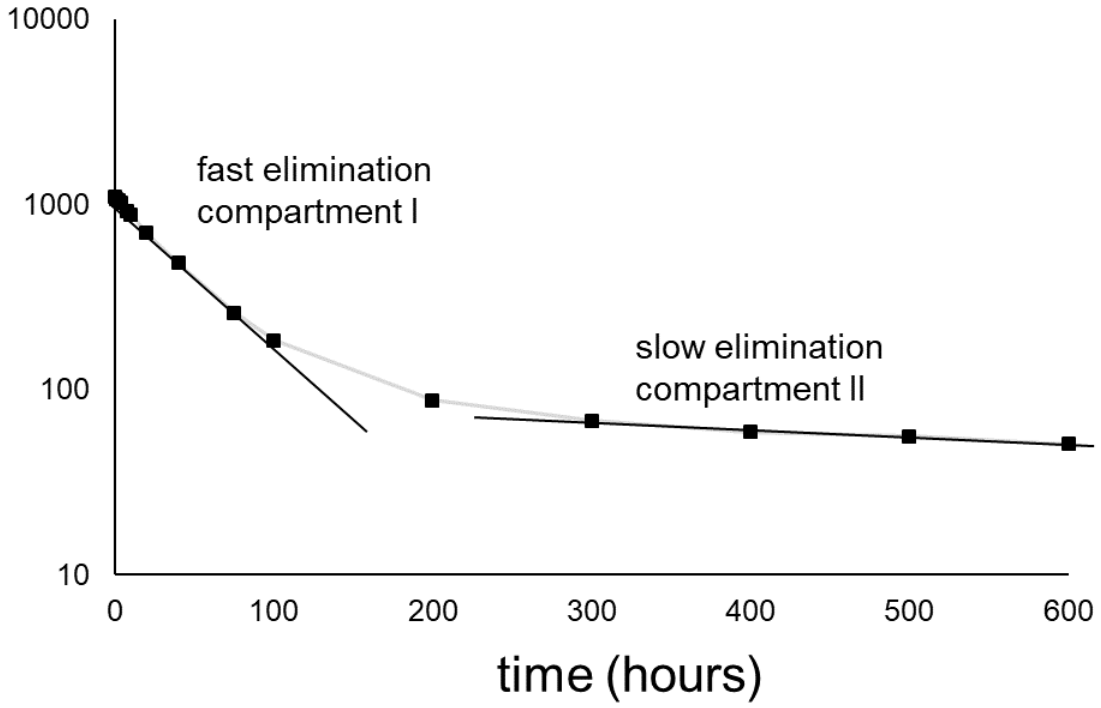


Figure 4.

Elimination of chemical from fish to water (top) in a two-compartment model (bottom).

Examples of uptake curves for different chemicals and organisms

Figure 5 gives uptake curves for two different chemicals and the corresponding kinetic parameters. Chemical 2 has a BCF of 1000, chemical 1 a BCF of 10,000. Uptake rates (k_w) are the same, which is often the case for organic chemicals. Half-lives (time to reach 50 % of the steady state level) are 14 and 140 hours. This makes sense because it will take a longer time to reach steady state for a chemical with a higher BCF. Elimination rate constants also differ a factor of 10.

In figure 6, uptake curves are presented for one chemical, but in two organisms of different size/weight. Organism 1 is much smaller than organism 2 and reaches steady state much earlier. $T_{1/2}$ values for the chemical in organisms 1 and 2 are 14 and 140 hours, respectively. The small size explains this fast equilibration. Rates of uptake depend on the surface-to-volume ratio (S/V) of an organism, which is much higher for a small organism. Therefore, kinetics in small organisms is faster resulting in shorter equilibration times. The effect of size on kinetics is discussed in more detail in Hendriks et al. (2001) and in the Section on Allometric Relationships.

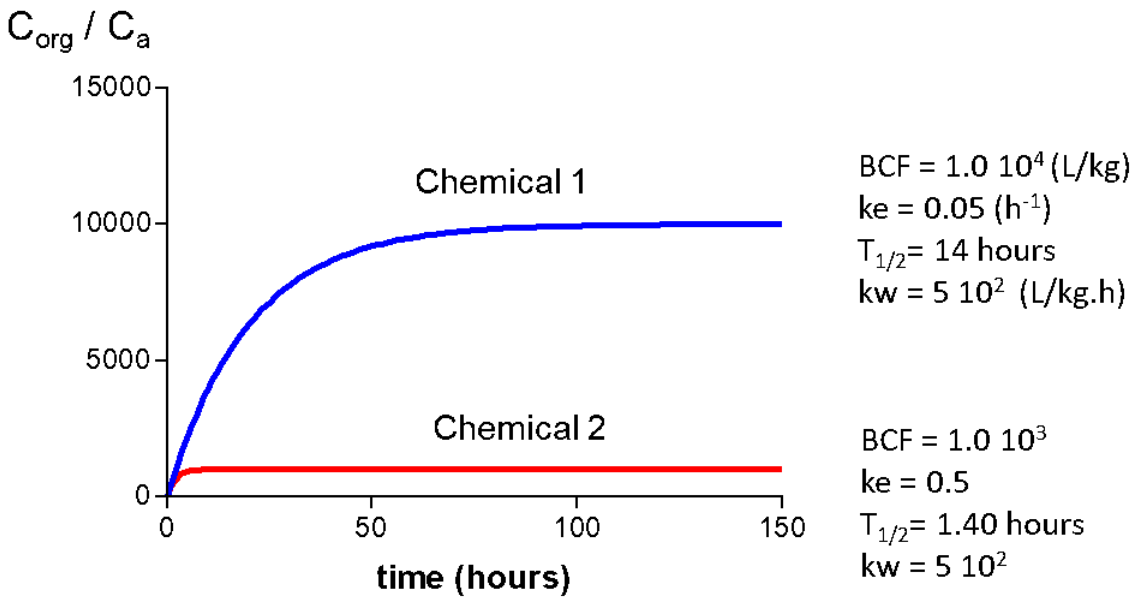


Figure 5. Uptake curves for two chemicals, having different properties, in the same organism.

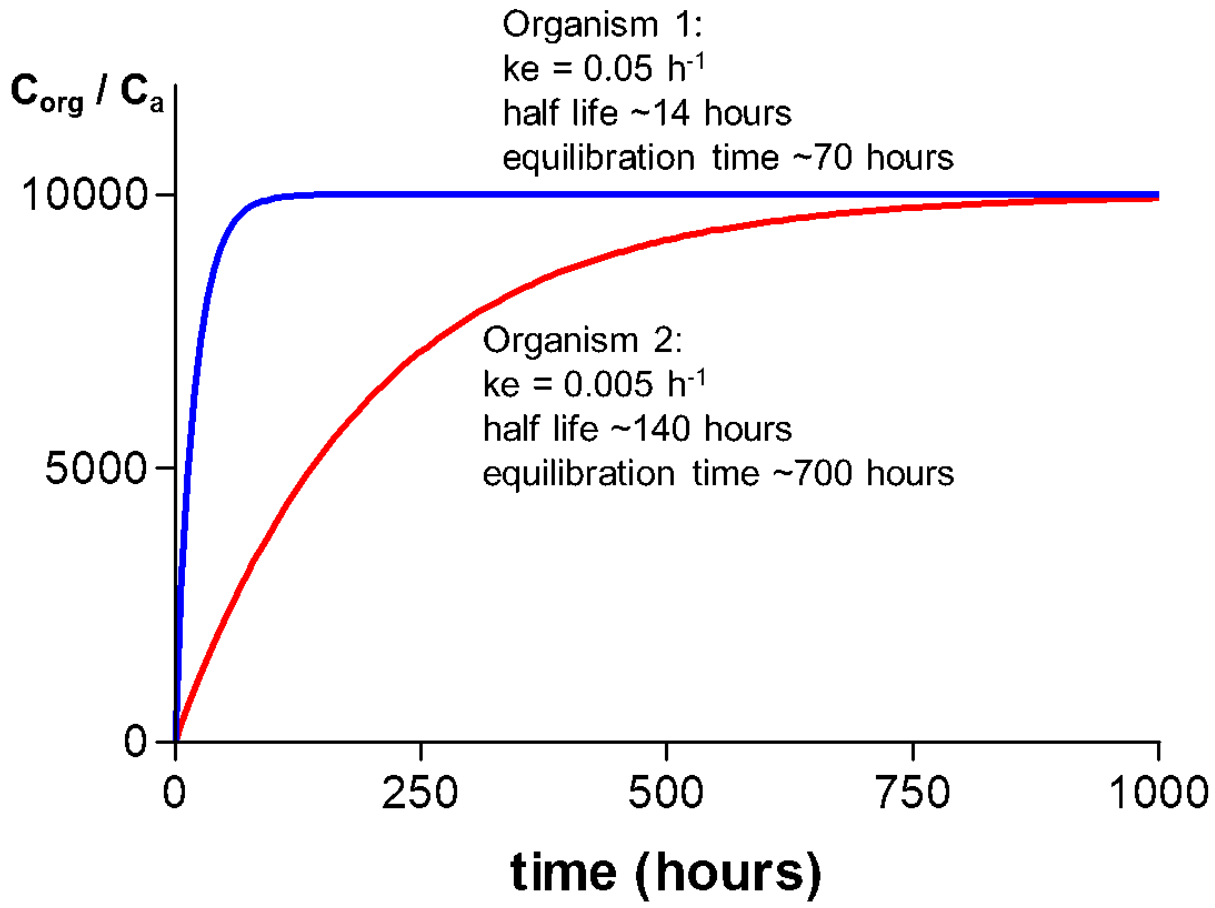


Figure 6. Uptake curves for the same chemical in two organisms having different sizes; organism 2 is much bigger than organism 1.

Bioaccumulation involving biotransformation and different routes of uptake

In equation 2, elimination only includes gill elimination. If other processes such as biotransformation and growth are considered, the equation can be extended to include these additional processes (see equation 12).

$$\frac{dC_{org}}{dt} = k_w C_a - (k_e + k_m + k_r) C_{org} \quad (12)$$

For organisms living in soil or sediment, different routes of uptake may be of importance: dermal (across the skin), or oral (by ingestion of food and/or soil or sediment particles). Mathematically, the uptake in an organism in sediment can be described as in equation 13.

$$C_{org}(t) = \frac{(k_w C_{aq} + k_s C_s)}{k_e}(1 - e^{-k_e t}) \quad (13)$$

C_{org} concentration in organisms (mg/kg)

C_{aq} concentration in aqueous phase (mg/L)

C_s concentration soil or sediment (mg/kg)

k_w uptake rate constant from water (L/kg/day)

k_s uptake rate constant from soil or sediment ($kg_{soil}/kg_{organism}$/day)

k_e elimination rate constant (1/day)

t time (day)

(dimensions used are: amount of chemical: mg; volume of water: L; weight of organism: kg; time: day)

In this equation, k_w and k_s are the uptake rate constants from water and sediment, k_e is the elimination rate constant and C_{aq} and C_s are the concentrations in water and sediment or soil. For soil organisms, such as earthworms, oral uptake appears to become more important with increasing hydrophobicity of the chemical (Jager et al., 2003). This is because the concentration in soil (C_s) will become higher than the porewater concentration C_a for the more hydrophobic chemicals (see section on Sorption).

References

Clewell, H.J., 3rd (1995). The application of physiologically based pharmacokinetic modeling in human health risk assessment of hazardous substances. Toxicology Letters 79, 207-217.

Hendriks, A.J., van der Linde, A., Cornelissen, G., Sijm, D. (2001). The power of size. 1. Rate constants and equilibrium ratios for accumulation of organic substances related to octanol-water partition ratio and species weight. *Environmental Toxicology and Chemistry* 20, 1399-1420.

Jager, T., Fleuren, R., Hogendoorn, E.A., De Korte, G. (2003). Elucidating the routes of exposure for organic chemicals in the earthworm, *Eisenia andrei (*Oligochaeta). Environmental Science and Technology 37, 3399-3404.

Nichols, J.W., Fitzsimmons, P.N., Whiteman, F.W. (2004). A physiologically based toxicokinetic model for dietary uptake of hydrophobic organic compounds by fish - II. Simulation of chronic exposure scenarios. Toxicological Sciences 77, 219-229.

Van Leeuwen, C.J., Vermeire, T.G. (Eds.) (2007). Risk Assessment of Chemicals: An Introduction. Springer, Dordrecht, The Netherlands.

4.1.2. Question 1

What are the assumptions in a one-compartment kinetic model?

4.1.2. Question 2

How can you identify when more compartments are involved in a kinetic model?

4.1.2. Question 3

Give examples of compartments in a multi-compartment system.

4.1.2. Question 4

Why is equilibrium reached faster in a small organism compared to a bigger organism?

4.1.2. Question 5

Which two methods can be applied to estimate the bioconcentration factor of a chemical in an organism?

4.1.3. Tissue accumulation of metals

Author: Nico M. van Straalen

Reviewers: Philip S. Rainbow, Henk Schat

Learning objectives:

You should be able to

- indicate four types of inorganic metal binding cellular constituents present in biological tissues and indicate which metals they bind.
- describe how phytochelatin and metallothionein are induced by metals.
- mention a number of organ-metal combinations that are critical to metal toxicity.

Key words: Metal binding proteins; phytochelatin; metallothionein;

Synopsis

The issue of metal speciation, which is crucially important to understand metal fate in the environment, is equally important for internal distribution in organisms and toxicity inside the cell. Many metals tend to accumulate in specific organs, for example the hepatopancreas of crustaceans, the chloragogen tissue of annelids, and the kidney of mammals. In addition, there is often a specific organ or tissue where dysfunction or toxicity is first observed, e.g., in the human body, the primary effects of chronic exposure to mercury are seen in the brain, for lead in bone marrow and for cadmium in the kidney. This module is aimed at increasing the insight into the different mechanisms by which metals accumulate in biological tissues.

Introduction

Metals will be present in biological tissues in a large variety of chemical forms: free metal ion, various inorganic species with widely varying solubility, such as chlorides or carbonates, plus all kind of metal species bound to low-molecular and high-molecular weight biotic ligands. The free metal ion is considered the most relevant species relating to toxicity.

To explain the affinities of metals with specific targets, a system has been proposed based on the physical properties of the ion; according to this system, metals are divided into "oxygen-seeking metals" (class A, e.g. lithium, beryllium, calcium and lanthanum), and "sulfur-seeking metals" (class B, e.g. silver, mercury and lead) (See section on Metals and metalloids). However, most of the metals of environmental relevance fall in an intermediate class, called "borderline" (chromium, cadmium, copper, zinc, etc.). This classification is to some extent predictive of the binding of metals to specific cellular targets, such as SH-groups in proteins, nitrogen in histidine or carbonates in bone tissue.

Not only do metals differ enormously in their physicochemical properties, also the organisms themselves differ widely in the way they deal with metals. The type of ligand to which a metal is bound, and how this ligand is transported or stored in the body, determines to a great extent where the metal will accumulate and cause toxicity. Sensitive targets or critical biochemical processes differ between species and this may also lead to differential toxicity.

Inorganic metal binding

Many tissues contain "mineral concretions", that is, granules with a specific mineral composition that, due to the nature of the mineral, attract different metals. Especially the gut epithelium of invertebrates, and their digestive glands (hepatopancreas, midgut gland, chloragogen tissue) may be full of such concretions. Four classes of granules are distinguished (Figure 1):

- Calcium-pyrophosphate granules with magnesium, manganese, often also zinc, cadmium, lead and iron
- Sulfur granules with copper, sometimes also cadmium
- Iron granules, containing exclusively iron
- Calcium carbonate granules, containing mostly calcium

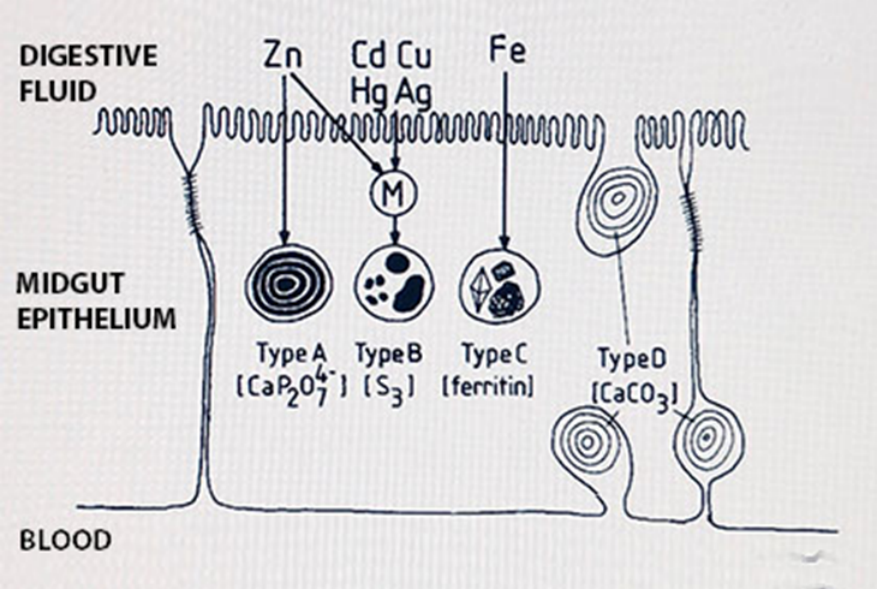


Figure 1. Schematic diagram of the gut epithelium of an invertebrate animal, showing four different types of granules and the metals they usually contain. Redrawn from Hopkin (1989) by Wilma IJzerman.

The type B granules are assumed to be lysosomal vesicles that have absorbed metal-loaded peptides such as metallothionein or phytochelatin, and have developed into inorganic granules by degrading almost all organic material; the high sulfur content derives from the cysteine residues in the peptides.

Tissues or cells that specialize in the synthesis of intracellular granules are also the places where metals tend to accumulate. Well-known are the "S cells" in the hepatopancreas of isopods. These cells (small cells, B-type cells *sensu* Hopkin 1989) contain very large amounts of copper. Most likely the large stores of copper in woodlice and other crustaceans relate to their use of hemocyanin, a copper-dependent protein, as an oxygen-transporting molecule. Similar tissues with high loadings of mineral concretions have been described for earthworms, snails, collembolans and insects.

Organic metal binding

The second class of metal-binding ligands is of organic nature. Many plants but also several animals synthesize a peptide called phytochelatin (PC). This is an oligomer derived from glutathione with the three amino acids, γ-glumatic acid, cysteine and glycine, arranged in the following way: $(\gamma\text{-glu-cys})_n$-gly, where n can vary from 2 to 11. The thiol groups of several cysteine residues are involved in metal binding.

The other main organic ligand for metals is metallothionein (MT). This is a low-molecular weight protein with hydrophilic properties and an unusually large number of cysteine residues. Several cysteines (usually nine or ten) can bind a number of metal ions (e.g. four or five) in one cluster. There are two such clusters in the vertebrate metallothionein. Metallothioneins occur throughout the tree of life, from bacteria to mammals, but the amino acid sequence, domain structure and metal affinities vary enormously and it is doubtful whether they represent a single evolutionary-homologous group.

In addition to these two specific classes of organic ligands, MT and PC, metals will also bind aspecifically to all kind of cellular constituents, such as cell wall components, albumen in the blood, etc. Often this represents the largest store of metals; such aspecific binding sites will constantly deliver free metal ions to the cellular pool and so are the most important cause of toxicity. Of course metals are also present in molecules with specific metal-dependent functions, such as iron in hemoglobin, copper in hemocyanin, zinc in carbonic anhydrase, etc.

The distinction between inorganic ligands and organic ones is not as strict as it may seem. After binding to metallothionein or phytochelatin metals may be transferred to a more permanent storage compartment, such as the intracellular granules mentioned above, or they may be excreted.

Regulation of metal binding

Free metal ions are strong inducers of stress response pathways. This can be due to the metal ion itself but more often the stress response is triggered by a metal-induced disturbance of the redox state, i.e. an induction of oxidative stress. The stress response often involves the synthesis of metal-binding ligands such as phytochelatin and metallothionein. Because this removes metal ions from the active pool it is also called metal scavenging.

The binding capacity of phytochelatin is enhanced by activation of the enzyme phytochelatin synthase (PC synthase). According to one model of its action, the C-terminus of the enzyme has a "metal sensor" consisting of a number of cysteines with free SH-groups. Any metal ions reacting with this nucleophile center (and cadmium is a strong reactant) will activate the enzyme which then catalyzes the reaction from $(\gamma\text{-glu-cys})_n\text{-gly}$ to $(\gamma\text{-glu-cys})_{n+1}\text{-gly}$, thus increasing the binding capacity of cellular phytochelatin (Figure 2). This reaction of course relies on the presence of sufficient glutathione in the cell. In plants the PC-metal complex is transported into the central vacuole, where it can be stabilized through incorporation of acid-labile sulfur (S^2). The PC moiety is degraded, resulting in the formation of inorganic metal-sulfide crystallites. Alternatively, complexes of metals with organic acids may be formed (e.g. citrates or oxalates). The fate of metal-loaded PC in animal cells is not known, but it might be absorbed in the lysosomal system to form B-type granules (see above).

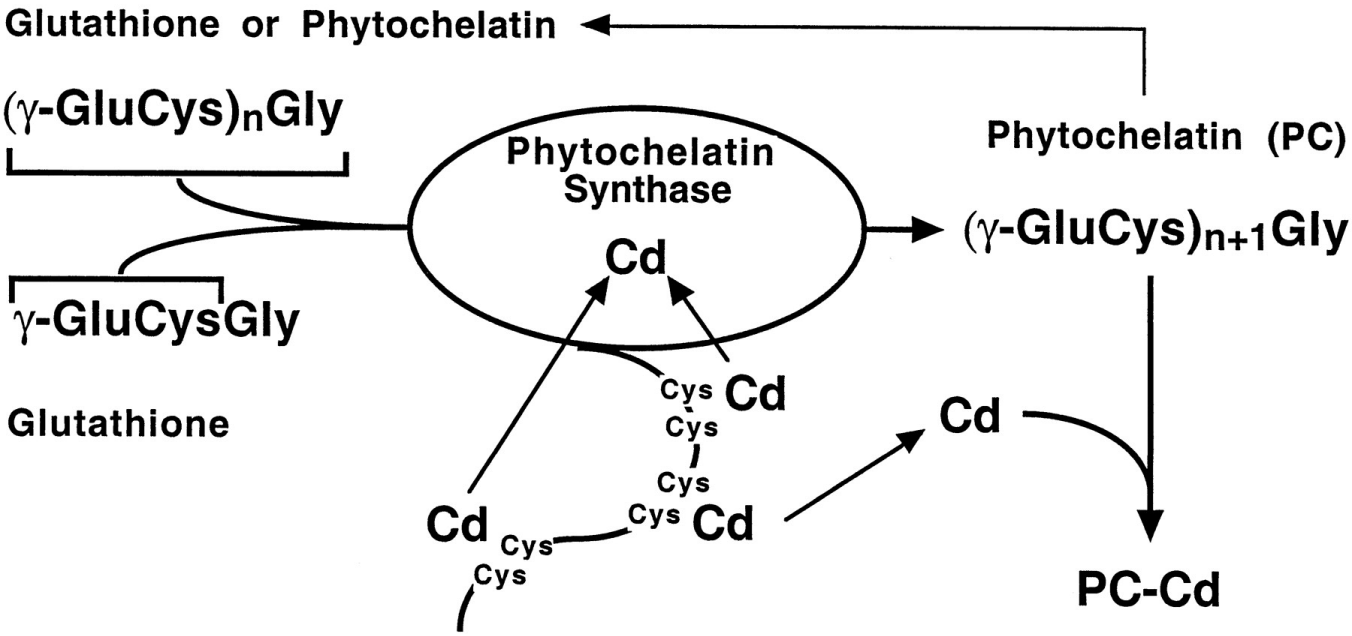


Figure 2. Model for the regulation of phytochelatin synthase, an enzyme catalyzing the extension of phytochelatin and increasing the binding capacity for metals. From Cobbet (2000). Source: http://www.plantphysiol.org/content/123/3/825

The upregulation of metallothionein (MT) occurs in a quite different manner, since it depends on *de novo* synthesis of the apoprotein. It is a classic example of gene regulation contributing to protection of the cell. In a wide variety of animals, including vertebrates and invertebrates, metallothionein genes (*Mt*) are activated by a transcription factor called metal-responsive transcription factor 1 (MTF-1). MTF -1 binds to so-called metal-responsive elements (MREs) in the promoter of *Mt*. MREs are short motives with a characteristic base-pair sequence that form the core of a transcription factor binding site in the DNA. Under normal physiological conditions MTF-1 is inactive and unable to induce *Mt*. However, it may be activated by Zn^{2+} ions, which are released, from unspecified ligands, by metals such as cadmium that can replace zinc (Figure 3).

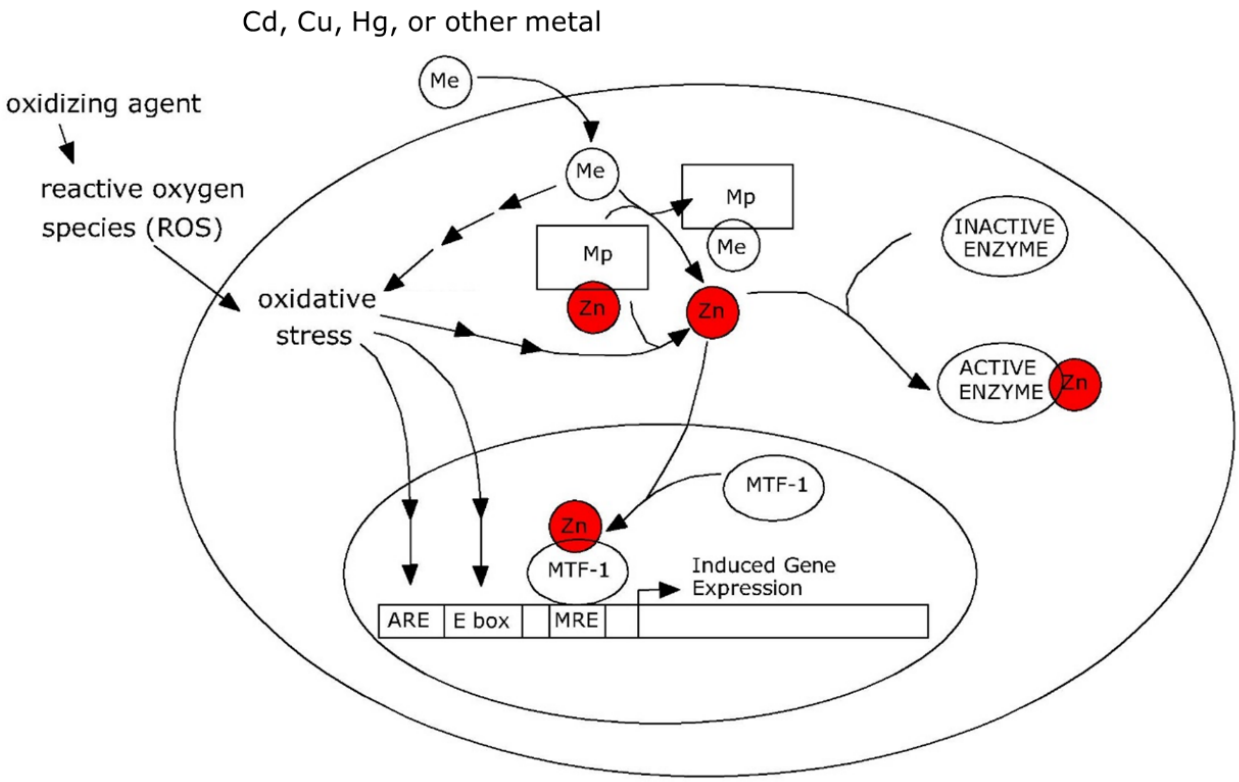


Figure 3. Model for metallothionein induction by cadmium. Cadmium ions (Me) displace zinc ions from ligands (MP). Free zinc ions (Zn) then activate Metal Transcription Factor 1 (MTF-1) in the nucleus to bind to transcription factor binding sites containing Metal-Responsive Elements (MRE). This induces expression of Mt and other genes. These genes may also be induced by oxidative stress, acting upon antioxidant-responsive elements (ARE) and Enhancer Box (E box). Oxidative stress may come from external oxidizing agents or from the metal itself. Finally, zinc ions may also directly activate antioxidant enzymes. Adapted from Haq et al. (2003) by Wilma IJzerman.

It must be emphasized that the model discussed above is inspired by work on vertebrates. Arthropods (*Drosophila*, *Orchesella*, *Daphnia*) could have a similar mechanism since they also have an MTF-1 homolog that activates *Mt*, however, the situation for other invertebrates such as annelids and gastropods is unclear; their *Mt* genes seem to lack MREs, despite being inducible by cadmium. In addition, the variability of metallothioneins in invertebrates is extremely large and not all metal-binding proteins may be orthologs of the vertebrate metallothionein. In snails, a cadmium-binding, cadmium-induced MT functions alongside a copper-binding MT while the two MTs have different tissue distributions and are also regulated quite differently.

While both phytochelatin and metallothionein will sequester essential as well as non-essential metals (e.g. Cd) and so contribute to detoxification, the widespread presence of these systems throughout the tree of life suggests that they did not evolve primarily to deal with anthropogenic metal pollution. The very strong inducibility of these systems by non-essential elements like cadmium may be considered a side-effect of a different primary function, for example regulation of the cellular redox state or binding of essential metals.

Target organs

Any tissue-specific accumulation of metals can be explained by turnover of metal-binding ligands. For example, accumulation of cadmium in mammalian kidney is due to the fact that metallothionein loaded with cadmium cannot be excreted. High concentrations of metals in the hind segments of earthworms are due to the presence of "residual bodies" which are fully packed with intracellular granules. Accumulation of cadmium in the human prostrate is due to the high concentration of zinc citrate in this organ, which serves to protect contractile proteins in sperm tails from oxidation; cadmium assumedly enters the prostrate through zinc transporters.

It is often stated that essential metals are subject to regulatory mechanisms, which would imply that their body burden, over a large range of external exposures, is constant. However, not all "essential" metals are regulated to the extent that the whole-body

concentration is kept constant. Many invertebrates have body compartments associated with the gut (midgut gland, hepatopancreas, Malpighian tubules) in which metals, often in the form of mineral concretions, are inactivated and stored permanently or exchanged very slowly with the active pool. Since these compartments are outside the reach of regulatory mechanisms but usually not separated in whole-body metal analysis, the body burden as a whole is not constant. Some invertebrates even carry a "backpack" of metals accumulating over life. This holds, e.g., for zinc in barnacles, copper in isopods and zinc in earthworms.

Accumulation of metals in target organs may lead to toxicity when the critical binding or excretion capacity is exhausted and metal ions start binding aspecifically to cellular constituents. The organ in which this happens is often called the target organ. The total metal concentration at which toxicity starts to become apparent is called the critical body concentration (CBC) or critical tissue concentration. For example, the critical concentration for cadmium in kidney, above which renal damage is observed to occur, is estimated to be 50 µg/g. A list of critical organs for metals in the human body is given in Table 1.

The concept of CBC assumes that the complete metal load in an organ is in equilibrium with the active fraction causing toxicity and that there is no permanent storage pool. In the case of storage detoxification the body burden at which toxicity appears will depend on the accumulation history.

Table 1. Critical organs for chronic toxicity of metals in the human body

Metal or metalloid	Critical organ	Symptoms
Al	Brain	Alzheimer's disease
As	Lung, liver, heart gut	Multisystem energy disturbance
Cd	Kidney, liver	Kidney damage
Cr	Skin, lung, gut	Respiratory system damage
Cu	Liver	Liver damage
Hg	Brain, liver	Mental illness
Ni	Skin, kidney	Allergic reaction, kidney damage
Pb	Bone marrow, blood, brain	Anemia, mental retardation

References

Cobbett, C., Goldsbrough, P. (2002). Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. Annual Review of Plant Biology 53, 159-182.

Dallinger, R., Berger, B., Hunziker, P., Kägi, J.H.R. (1997). Metallothionein in snail Cd and Cu metabolism. Nature 388, 237-238.

Dallinger, R., Höckner, M. (2013). Evolutionary concepts in ecotoxicology: tracing the genetic background of differential cadmium sensitivities in invertebrate lineages. Ecotoxicology 22, 767-778.

Haq, F., Mahoney, M., Koropatnick, J. (2003) Signaling events for metallothionein induction. Mutation Research 533, 211-226.

Hopkin, S.P. (1989) *Ecophysiology of Metals in Terrestrial Invertebrates*. London, Elsevier Applied Science.

Nieboer, E., Richardson, D.H.S. (1980) The replacement of the nondescript term "heavy metals" by a biologically and chemically significant classification of metal ions. Environmental Pollution Series B 1, 3-26.

Rainbow, P.S. (2002) Trace metal concentrations in aquatic invertebrates: why and so what? Environmental Pollution 120, 497-507.

4.1.3. Question 1

Mention the four main inorganic cellular constituents that bind metals, their main elemental composition and indicate what metals are usually bound in each structure.

4.1.3. Question 2

How can the intracellular levels of metal-scavenging molecules such as metallothionein (MT) and phytochelatin (PC) be adjusted to counteract the possible adverse effects of free metal ions? Describe the molecular mechanism for metallothionein and phytochelatin separately.

4.1.3. Question 3

1. Mention three organs of the human body susceptible to metal intoxication and to which metals they are particularly sensitive.

4.1.4. Xenobiotic defence and metabolism

Author: Nico M. van Straalen

Reviewers: Timo Hamers, Cristina Fossi

Learning objectives:

You should be able to:

- recapitulate the phase I, II and III mechanisms for xenobiotic metabolism, and the most important molecular systems involved.
- describe the fate and the chemical changes of an organic compound that is metabolized by the human body, from absorption to excretion.
- explain the principle of metabolic activation and why some compounds can become very reactive upon xenobiotic metabolism.
- develop a hypothesis on the ecological effects of xenobiotic compounds that require metabolic activation.

Keywords: biotransformation; phase 1; phase II; phase III; excretion; metabolic activation; cytochrome P450

Synopsis

All organisms are equipped with metabolic defence mechanisms to deal with foreign compounds. The reactions involved, jointly called biotransformation, can be divided into three phases, and usually aim to increase water solubility and excretion. The first step (phase I) is catalyzed by cytochrome P450, which is followed by a variety of conjugation reactions (phase II) and excretion (phase III). The enzymes and transporters involved are often highly inducible, i.e. the amount of protein is greatly enhanced by the xenobiotic compounds themselves. The induction involves binding of the compound to cytoplasmic receptor proteins, such as the arylhydrocarbon receptor (AhR), or the constitutive androstane receptor (CAR). In some cases the intermediate metabolites, produced in phase I are extremely reactive and a main cause of toxicity, a well-known example being the metabolic activation of polycyclic aromatic hydrocarbons such as benzo(a)pyrene, which readily forms DNA adducts and causes cancer. In addition, some compounds greatly induce metabolizing enzymes but are hardly degraded by them and cause chronic cellular stress. The various biotransformation reactions are a crucial aspect of both toxicokinetics and toxicodynamics of xenobiotics.

Introduction

The term "xenobiotic" ("foreign to biology") is generally used to indicate a chemical compound that does not normally have a metabolic function. We will use the term extensively in this module, despite the fact that it is somewhat problematic (e.g., can a compound be considered "foreign" if it circulates in the body, is metabolized or degraded in the body?, and: what is "foreign" for one species is not necessarily "foreign" for another species).

The body has an extensive defence system to deal with xenobiotic compounds, loosely designated as biotransformation. The ultimate result of this system is excretion the compound in some form or another. However, many xenobiotics are quite lipophilic, tend to accumulate and are not easily excreted due to low water solubility. Molecular modifications are usually required before such compounds can be removed from the body, as the main circulatory and excretory systems (blood, urine) are water-based. By introducing hydrophilic groups in the molecule (-OH, =O, -COOH) and by conjugating it to an endogenous compound with good water-solubility, excretion is usually accomplished. However, as we will see below, intermediate metabolites may have enhanced reactivity and it often happens that a compound becomes more toxic while being metabolized. In the case of pesticides, deliberate use is made of such responses, to increase the toxicity of an insecticide once it is in the target organism.

The study of xenobiotic metabolism is a classical subject not only in toxicology but also in pharmacology. The mode of action of a drug often depends critically on the rate and mode of metabolism. Also, many drugs show toxic side-effects as a consequence of

metabolism. Finally, xenobiotic metabolism is also studied extensively in entomology, as both toxicity and resistance of pesticides are often mediated by metabolism.

The most problematic xenobiotics are those with a high octanol-water partition coefficient (K_{ow}) that are strongly lipophilic and very hydrophobic. They tend to accumulate, in proportion to their Log K_{ow}, in tissues with a high lipid content such as the subcutis of vertebrates, and may cause tissue damage due to disturbance of membrane functions. This mode of action is called "minimum toxicity". Well-known are low-molecular weight aliphatic petroleum compounds and chlorinated alkanes such as chloroform. These compounds cause their primary damage to cell membranes; especially neurons are sensitive to this effect, hence minimum toxicity is also called narcotic toxicity. Lipophilic chemicals with high Log K_{ow} do not reach concentrations high enough to cause minimum toxicity because they induce biotransformation at lower concentrations. The toxicity is then usually due to a reactive metabolite.

Xenobiotic metabolism involves three subsequent phases (Figure 1):

1. Activation (usually oxidation) of the compound by an enzyme known as cytochrome P450, which acts in cooperation with NADPH cytochrome P450 reductase and other factors.
2. Conjugation of the activated product of phase I to an endogenous compound. A host of different enzymes is available for this task, depending on the compound, the tissue and the species. There are also (slightly polar) compounds that enter phase II directly, without being activated in phase I.
3. Excretion of the compound into circulation, urine, or other media, usually by means of membrane-spanning transporters belonging to the class of ATP-binding cassette (ABC) transporters, including the infamous multidrug resistance proteins. Hydrophilic compounds may pass on directly to phase III, without being activated or conjugated.

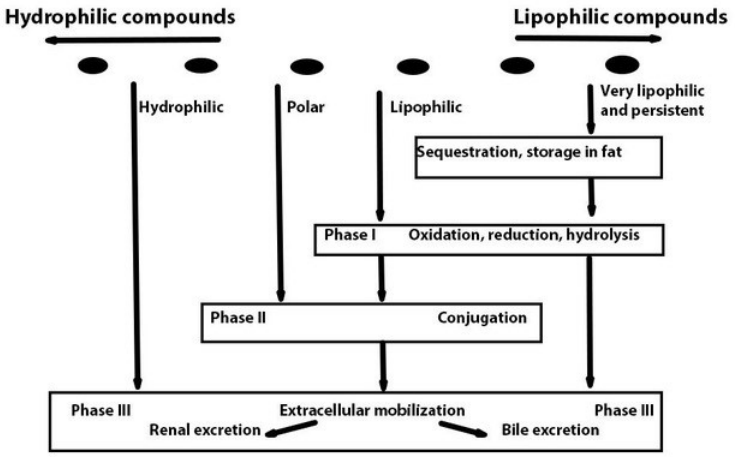


Figure 1. The various phases of xenobiotic metabolism. Redrawn by Wilma Ijzerman.

Phase I reactions

Cytochrome P450 is a membrane-bound enzyme, associated with the smooth endoplasmic reticulum. It carries a porphyrin ring containing an Fe atom, which is the active center of the molecule. The designation P450 is derived from the fact that it shows an absorption maximum at 450 nm when inhibited by carbon monoxide, a now outdated method to demonstrate its presence. Other (outdated) namings are MFO (mixed function oxygenase) and drug metabolizing enzyme complex. Cytochrome P450 is encoded by a gene called CYP, of which there are many paralogs in the genome, all slightly differing from each other in terms of inducibility and substrate specificity. Three classes of CYP genes are involved with biotransformation, designated CYP1, CYP2 and CYP3 in vertebrates. Each class has several isoforms; the human genome has 57 different CYP genes in total. The CYP complement of invertebrates and plants often involves even more genes; many evolutionary lineages have their own set, arising from extensive gene duplications within that lineage. In humans, the genetic complement of a person's CYP genes is highly relevant as to its drug metabolizing profile (see the section on Genetic variation in toxicant metabolism).

Cytochrome P450 operates in conjunction with an enzyme called NADPH cytochrome P450 reductase, which consists of two flavoproteins, one containing Flavin adenine dinucleotide (FAD), the other flavin mononucleotide (FMN). The reduced Fe^{2+} atom in cytochrome P450 binds molecular oxygen, and is oxidized to Fe^{3+} while splitting O_2; one O atom is introduced in the substrate,

the other reacts with hydrogen to form water. Then the enzyme is reduced by accepting an electron from cytochrome P450 reductase. The overall reaction can be written as:

$RH + O_2 + NADPH + H^+ \rightarrow ROH + H_2O + NADP^+$

where R is an arbitrary substrate.

Cytochrome P450 is expressed to a great extent in hepatocytes (liver cells), the liver being the main organ for xenobiotic metabolism in vertebrates (Figure 2), but it is also present in epithelia of the lung and the intestine. In insects the activity is particularly high in the Malpighian tubules in addition to the gut and the fat body. In mollusks and crustaceans the main metabolic organ is the hepatopancreas.

Phase II reactions

After activation by cytochrome P450 the oxidized substrate is ready to be conjugated to an endogenous compound, e.g. a sulphate, glucose, glucuronic acid or glutathione group. These reactions are conducted by a variety of different enzymes, of which some reside in the sER like P450, while others are located in the cytoplasm of the cell (Figure 2). Most of them transfer a hydrophilic group, available from intermediate metabolism, to the substrate, hence the enzymes are called transferases. Usually the compound becomes more polar in phase II, however, not all phase II reactions increase water solubility; for example, methylation (by methyl transferase) decreases reactivity but increases apolarity. Other phase II reactions are conjugation with glutathione, conducted by glutathione-S-transferase (GST) and with glucuronic acid, conducted by UDP-glucuronyl transferase. In invertebrates other conjugations may dominate, e.g. in arthropods and plants, conjugation with malonyl glucose is a common reaction, which is not seen in vertebrates.

Conjugation with glutathione in the human body is often followed by splitting off glutamic acid and glycine, leaving only the cysteine residue on the substrate. Cysteine is subsequently acetylated, thus forming a so-called mercapturic acid. This is the most common type of metabolite for many xenobiotics excreted in urine by humans.

Like cytochrome P450, the phase II enzymes consist in various isoforms, encoded by different paralogs in the genome. Especially the GST family is quite extensive and polymorphisms in these genes contribute significantly to the personal metabolic profile (see the section on Genetic variation in toxicant metabolism).

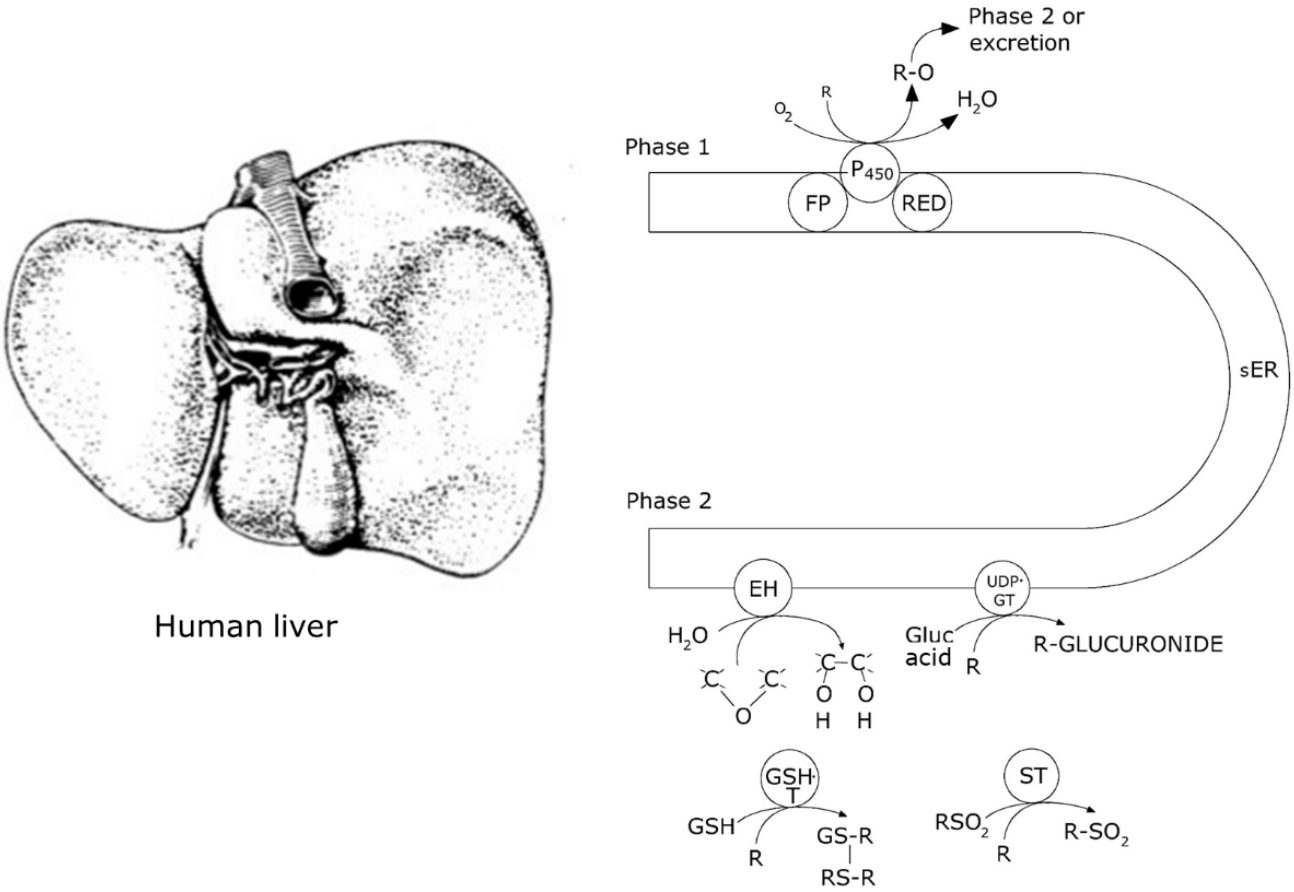


Figure 2. Schematic view of xenobiotic metabolism (phase I and phase II) in human liver. P450 = cytochrome P450, FP = flavoprotein, RED = cytochrome P450 reductase, EH = epoxide hydrolase, UDP-GT = uridyldiphospho-glucuronyl transferase, GSH-T = glutathione-S-transferase, ST = sulphotransferase. Reproduced from Vermeulen & Van den Broek (1984) by Wilma Ijzerman.

Phase III reactions

In the human body, there are two main pathways for excretion, one from the liver into the bile (and further into the gut and the faeces), the other through the kidney and urine. These two pathways are used by different classes of xenobiotics: very hydrophobic compounds such as high-molecular weight polycyclic aromatic hydrocarbons are still not readily soluble in water even after metabolism but can be emulsified by bile salt and excreted in this way. It sometimes happens that such compounds, once arriving in the gut, are assimilated again, transported to the liver by the portal vein and metabolized again. This is called "entero-hepatic circulation". Lower molecular weight compounds and hydrophilic compounds are excreted through urine. Volatile compounds can leave the body through the skin and exhaled air.

Excretion of activated and conjugated compounds from tissues out of the cell usually requires active transport, which is mediated by ABC (ATP binding cassette) transporters, a very large and diverse family of membrane proteins which have in common a binding cassette for ATP. Different subgroups of ATP transporters transport different types of chemicals, e.g. positively charged hydrophobic molecules, neutral molecules and water-soluble anionic compounds. One well-known group consists of multidrug resistance proteins or P-glycoproteins. These transporters export drugs aiming to attack tumor cells. Because their activity is highly inducible, these proteins can enhance excretion enormously, making the cell effectively resistant and thus case major problems for cancer treatment.

Induction

All enzymes of xenobiotic metabolism are highly inducible: their activity is normally on a low level but is greatly enhanced in the presence of xenobiotics. This is achieved through a classic case of transcriptional regulation upon CYP and other genes, leading to

de novo synthesis of protein. In addition, extensive proliferation of the endoplasmic reticulum may occur, and in extreme cases even swelling of the liver (hepatomegaly).

The best investigated pathway for transcriptional activation of CYP genes is due to the arylhydrocarbon receptor (AhR). Under normal conditions, this peptide is stabilized in the cytoplasm by heat-shock proteins, however, when a xenobiotic compound binds to AhR, it is activated and can join with another protein called Ah receptor nuclear translocator (ARNT) to translocate to the nucleus and bind to DNA elements present in the promotor of CYP and other genes. It thus acts as a transcriptional activator or transcription factor on these genes (Figure 3). The DNA motifs to which AhR binds are called xenobiotic responsive elements (XRE) or dioxin-responsive elements (DRE). The compounds acting in this manner are called 3-MC-type inducers, after the (highly carcinogenic) model compound 3-methylcholanthrene. The inducing capacity of a compound is related to its binding affinity to the AhR, which in itself is determined by the spatial structure of the molecule. The lock-and key fitting between AhR and xenobiotics explains why induction of biotransformation by xenobiotics shows a very strong stereospecificity. For example, among the chlorinated biphenyls and chlorinated dibenzodioxins, some compounds are extremely strong inducers of CYP1 genes, while others, even with the same number of chlorine atoms, are no inducers at all. The precise position of the chlorine atoms determines the molecular "fit" in the Ah receptor (see Section on Receptor interaction).

In addition to 3MC-type of induction there are other modes in which biotransformation enzymes are induced, but these are less well-known. A common class is PB-type induction (named after another model compound, phenobarbital). PB-type induction is not AhR-dependent, but acts through activation of another nuclear receptor, called constitutive androstane receptor (CAR). This receptor activates CYP2 genes and some CYP3 genes.

The high inducibility of biotransformation can be exploited in a reverse manner: if biotransformation is seen to be highly upregulated in a species living in the environment, this indicates that that species is being exposed to xenobiotic compounds. Assays addressing cytochrome P450 activity can therefore be exploited in bioindication and biomonitoring systems. The EROD (ethoxyresorufin-O-deethylase) assay is often used for this purpose, although it is not 100% specific to the isoform of P450. Another approach is to address CYP expression directly, e.g. through reverse transcription-quantitative PCR, a method to quantify the amount of CYP mRNA.

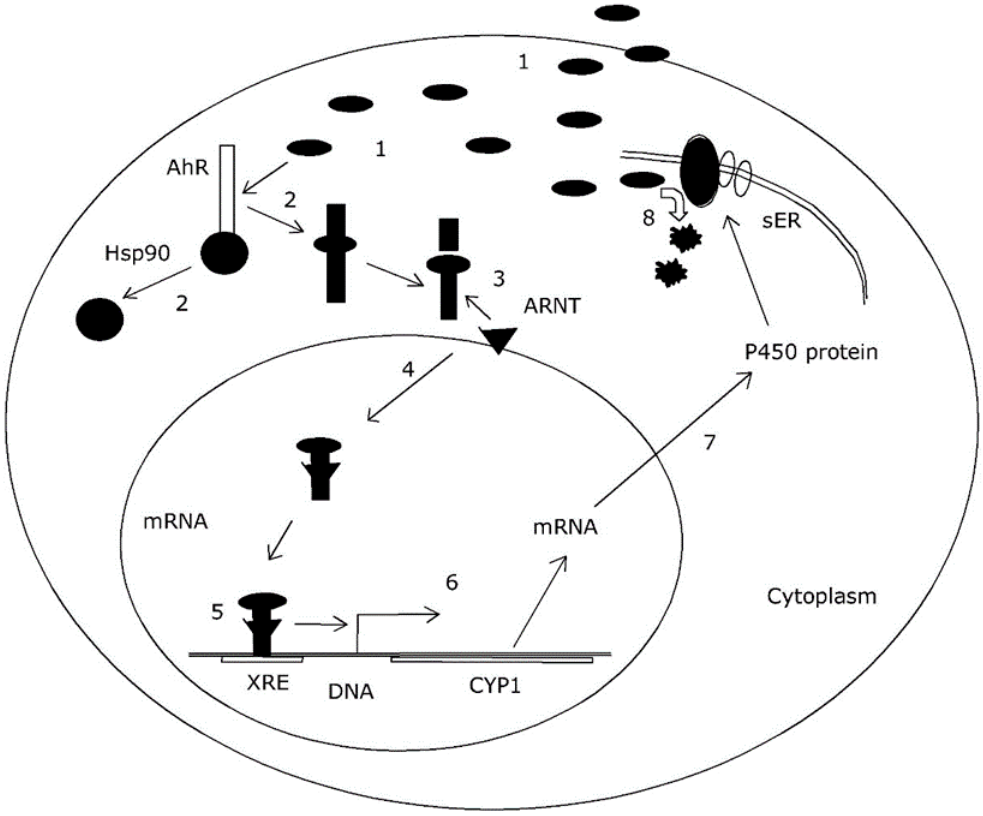


Figure 3. Scheme illustrating how AhR-reactive compounds induce biotransformation activity. 1 xenobiotic compounds enter the cell through

diffusion, 2 compound binds to AhR, releasing it from Hsp90, 3 AhR complex is activated by ARNT, 4 translocation to nucleus, 5 binding to XRE, 6 enhanced transcription, 7 translocation of mRNA to cytoplasm and translation, 8 incorporation of P450 enzyme in sER and biotransformation of compounds. AhR = arylhydrocarbon receptor, Hsp90 = heat shock protein 90, ARNT = aryl hydrocarbon nuclear translocator, XRE = xenobiotic responsive element, CYP1 = gene encoding cytochrome P450 1, sER = smooth endoplasmic reticulum. Drawn by Wilma Ijzerman.

Secondary effects of biotransformation

Although the main aim of xenobiotic metabolism is to detoxify and excrete foreign compounds, some pathways of biotransformation actually enhance toxicity. This is mostly due to the first step, activation by cytochrome P450. The activation may lead to intermediate metabolites which are highly reactive and the actual cause of toxicity. The best investigated examples are due to bioactivation of polycyclic aromatic hydrocarbons (PAHs), a group of chemicals present in diesel, soot, cigarette smoke and charred food products. Many of these compounds, e.g. benzo(a)pyrene, benz(a)anthracene and 3-methylcholanthrene, are not reactive or toxic as such but are activated by cytochrome P450 to extremely reactive molecules. Benzo(a)pyrene, for instance, is activated to a diol-epoxide, which readily binds to DNA, especially to the free amino-group of guanine (Figure 4). The complex is called a DNA adduct, the double helix is locally disrupted and this results in a mutation. If this happens in an oncogene, a tumor may develop (see the Section on Carcinogenesis and genotoxicity).

Not all PAHs are carcinogenic. Their activity critically depends on the spatial structure of the molecule, which again determines its "fit" in the Ah receptor. PAHs with a "notch" (often called bay-region) in the molecule tend to be stronger carcinogens than compounds with a symmetric (round or linear) molecular structure.

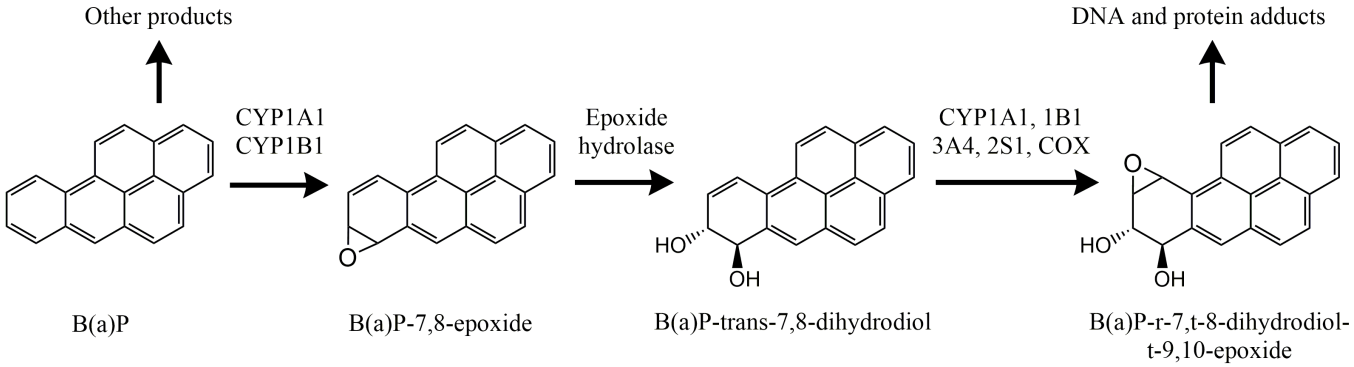


Figure 4. Metabolism of benzo(a)pyrene (B[a]P) to a mutagenic and carcinogenic metabolite, BaP-diol epoxide. Cytochrome P450 activity (CYP1A1 and CYP1B1) introduces an epoxide on the 7,8 position of the molecule, which is then hydrolyzed by epoxide hydrolase to a dihydrodiol (two OH groups next to each other, which can be in trans position, pointing in different directions relative to the plane of the molecule, or in cis, pointing in the same direction). The trans metabolite is preferably formed. Subsequently another epoxide is introduced on the 9,10 position. The diolepoxide is highly reactive, and may bind proteins and to guanine in DNA, causing a DNA adduct. Modified from Bui et al. (2009) by Steven Droge.

Another mechanism for biotransformation-induced toxicity is due to some very recalcitrant organochlorine compounds such as polychlorinated dibenzodioxins (PCDDs, or dioxins for short) and polychlorinated biphenyls (PCBs). Some of these compounds are very potent inducers of biotransformation, but they are hardly degraded themselves. The consequence is that the highly upregulated cytochrome P450 activity continues to generate a large amount of reactive oxygen (ROS), causing oxidative stress and damage to cellular constituents. It is assumed that the chronic toxicity of 2,3,7,8-tetrachlorodibenzo(para)dioxin (TCDD), one of the most toxic compounds emitted by human activity, is due to its high capacity to induce prolonged oxidative stress. On the molecular level, there is a close link between oxidative stress and biotransformation activity. Many toxicants that primarily induce oxidative stress (e.g. cadmium) also upregulate CYP enzymes. Two defence mechanisms, oxidative stress defence and biotransformation are part of the same integrated stress defence system of the cell.

References

Bui, P.H., Hsu, E.L., Hankinson, O. (2009), Fatty acid hydroperoxides support cytochrome P450 2S1-mediated bioactivation of benzo[a]pyrene-7-8-dihydrodiol. Molecular Pharmacology 76, 1044-1052.

Stroomberg, G.J., Zappey, H., Steen, R.J.C.A., Van Gestel, C.A.M., Ariese, F., Velthorst, N.H., Van Straalen, N.M. (2004). PAH biotransformation in terrestrial invertebrates - a new phase II metabolite in isopods and springtails. Comparative Biochemistry and Physiology Part C 138, 129-137.

Timbrell, J.A. (1982). *Principles of Biochemical Toxicology*. Taylor & Francis Ltd, London.

Van Straalen, N.M., Roelofs, D. (2012). *An Introduction to Ecological Genomics*, 2nd Ed. Oxford University Press, Oxford.

Vermeulen, N.P.E., Van den Broek, J.M. (1984). Opname en verwerking van chemicaliën in de mens. Chemisch Magazine. Maart: 167-171.

4.1.4. Question 1

Describe the main processes involved with each of the three general phases of xenobiotic metabolism, as well as the molecular systems involved.

4.1.4. Question 2

The upregulation of cytochrome P450 activity is a classical example of regulation through enhanced *de novo* synthesis of enzyme. The mechanism is known in quite some detail and involves a variety of components. Describe the involvement of the following components

- Aryl hydrocarbon receptor
- Heat shock protein 70
- Arylhydrocarbon receptor nuclear translocator
- Xenobiotic-response elements
- CYP1A1 gene
- CYP1A1 mRNA
- CYP 1A1 protein
- CYP 1A1 enzymatic activity

4.1.4. Question 3

In the past, activity of cytochrome P450 was often assessed using a synthetic fluorescent substrate, resorufin; activity was measured as ethoxyresorufin-O-deethylase (EROD). Discuss the advantages and disadvantages of the use of such an assay in environmental risk assessment.

4.1.5. Allometric relationships

Author: A. Jan Hendriks

Reviewers: Nico van den Brink, Nico van Straalen

Learning objectives:

You should be able to

- explain why allometrics is important in risk assessment across chemicals and species
- summarize how biological characteristics such as consumption, lifespan and abundance scale size
- describe how toxicological quantities such as uptake rates and lethal concentrations scale to size

Keywords: body size, biological properties, scaling, cross-species extrapolation, size-related uptake kinetics

Introduction

Globally more than 100,000,000 chemicals have been registered. In the European Union more than 100,000 compounds are awaiting risk assessment to protect ecosystem and human health, while 1,500,000 contaminated sites potentially require clean-up. Likewise, 8,000,000 species, of which 10,000 are endangered, need protection worldwide, with one lost per hour (Hendriks, 2013). Because of financial, practical and ethical (animal welfare) constraints, empirical studies alone cannot cover so many substances and species, let alone their combinations. Consequently, the traditional approach of ecotoxicological testing is gradually supplemented or replaced by modelling approaches. Environmental chemists and toxicologists for long have developed

relationships allowing extrapolation across chemicals. Nowadays, so-called Quantitative Structure Activity Relationships (QSARs) provide accumulation and toxicity estimates for compounds based on their physical-chemical properties. For instance bioaccumulation factors and median lethal concentrations have been related to molecular size and octanol-water partitioning, characteristic properties of a chemical that are usually available from its industrial production process.

In analogy with the QSAR approach in environmental chemistry, the question may be asked whether it is possible to predict toxicological, physiological and ecological characteristics of species from biological traits, especially traits that are easily measured, such as body size. This approach has gone under the name "Quantitative Species Sensitivity Relationships" (QSSR) (Notenboom, 1995).

Among the various traits available, body-size is of particular interest. It is easily measured and a large part of the variability between organisms can be explained from body size, with $r^2 > 0.5$. Not surprisingly, body size also plays an important role in toxicology and pharmacology. For instance, toxic endpoints, such as LC_{50}s, are often expressed per kg body weight. Recommended daily intake values assume a "standard" body weight, often 60 kg. Yet, adult humans can differ in body weight by a factor of 3 and the difference between mouse and human is even larger. Here it will be explored how body-size relationships, which have been studied in comparative biology for a long time, affect the extrapolation in toxicology and can be used to extrapolate between species.

Fundamentals of scaling in biology

Do you expect a 10^4 kg elephant to eat 10^4 times more than a 1 kg rabbit per day? Or less, or more? On being asked, most people intuitively come up with the right answer. Indeed, daily consumption by the proboscid is less than 10^4 times that of the rodent. Consequently, the amount of food or water used per kilogram of body weight of the elephant is less than that of the rabbit. Yet, how much less exactly? And why should sustaining 1 kg of rabbit tissue require more energy than 1 kg of elephant flesh in the first place?

A century of research (Peters, 1983) has demonstrated that many biological characteristics Y scale to size X according to a power function:

$$Y = a\,X^b$$

where the independent variable X represents body mass, and the dependent variable Y can virtually be any characteristic of interest ranging, e.g., from gill area of fish to density of insects in a community.

Plotted in a graph, the equation produces a curved line, increasing super-linearly if $b > 1$ and sub-linearly if $b < 1$. If b=1, Y and X are directly proportional and the relationship is called *isometric*. As curved lines are difficult to interpret, the equation is often simplified by taking the logarithm of the left and right parts. The formula then becomes:

$$\log Y = \log a + b \log X$$

When log Y is plotted against log X, a straight line results with slope b and intercept log a. If data are plotted in this way, the slope parameter b may be estimated by simple linear regression.

Across wide size ranges, slope b often turns out to be a multitude of ¼ or, occasionally, ⅓. Rates [$kg \cdot d^{-1}$] of consumption, growth, reproduction, survival and what not, increase with mass to the power ¾ , while rate constants, sometimes called specific rates [$kg \cdot kg^{-1} \cdot d^{-1}$], decrease with mass to the power -¼. So, while the elephant is 10^4 kg heavier than the 1 kg rabbit, it eats only $(10^4)^{¾} = 10^3$ times more each day. Vice versa, 1 kg of proboscid apparently requires a consumption of $(10^4)^{-¼}$ $kg \cdot kg^{-1} \cdot d^{-1}$, i.e., 10 times less. Variables with a time dimension [d] like lifespan or predator-prey oscillation periods scale inversely to rate constants and thus change with body mass to the power ¼. So, an elephant becomes $(10^4)^{¼} = 10$ times older than a rabbit. Abundance, i.e., the number of individuals per surface area [m^{-2}] decreases with body mass to the power -¾. Areas, such as gill surface or home range, scale inversely to abundance, typically as body mass to the power ¾.

Now, why would sustaining 1 kg of elephant require 10 times less food than 1 kg of rabbit? Biologists, pharmacologists and toxicologists first attributed this difference to area-volume relationships. If objects of the same shape but different size are compared, the volume increases with length to the power 3 and the surface increases with length to the power 2. For a sphere with radius r, for example, area A and volume V increase as $A \sim r^2$ and $V \sim r^3$, so area scales to volume as $A \sim r^2 \sim (V^{⅓})^2 \sim V^{⅔}$. So, larger animals have relatively smaller surfaces, as long as the shape of the organism remains the same. Since many biological

processes, such as oxygen and food uptake or heat loss deal with surfaces, metabolism was, for long, thought to slow down like geometric structures, i.e., with multitudes of ⅓. Yet, empirical regressions, e.g. the "mouse-elephant curve" developed by Max Kleiber in the early 1930s show a universal slope of ¼ (Peters, 1983. This became known as the "Kleiber's law". While the data leave little doubt that this is the case, it is not at all clear why it should be ¼ and not ⅓. Several explanations for the ¼ slope have been proposed but the debate on the exact value as well as the underlying mechanism continues.

Application of scaling in toxicology

Since chemical substances are carried by flows of air and water, and inside the organism by sap and blood, toxicokinetics and toxicodynamics are also expected to scale to size. Indeed, data confirm that uptake and elimination rate constants decrease with size, with an exponent of about -¼ (Figure 1). Slopes vary around this value, the more so for regressions that cover small size ranges and physiologically different organisms. The intercept is determined by resistances in unstirred water layers and membranes through which the substances pass, as well as by delays in the flows by which they are carried. The resistances mainly depend on the affinity and molecular size of the chemicals, reflected by, e.g., the octanol-water partition coefficient K_{ow} for organic chemicals or atomic mass for metals. The upper boundary of the intercept is set by the delays imposed by consumption and, subsequently, egestion and excretion. The lower end is determined by growth dilution. Both uptake and elimination scale to mass with the same exponent so that their ratio, reflecting the bioconcentration or biomagnification factor in equilibrium, is independent of body-size.

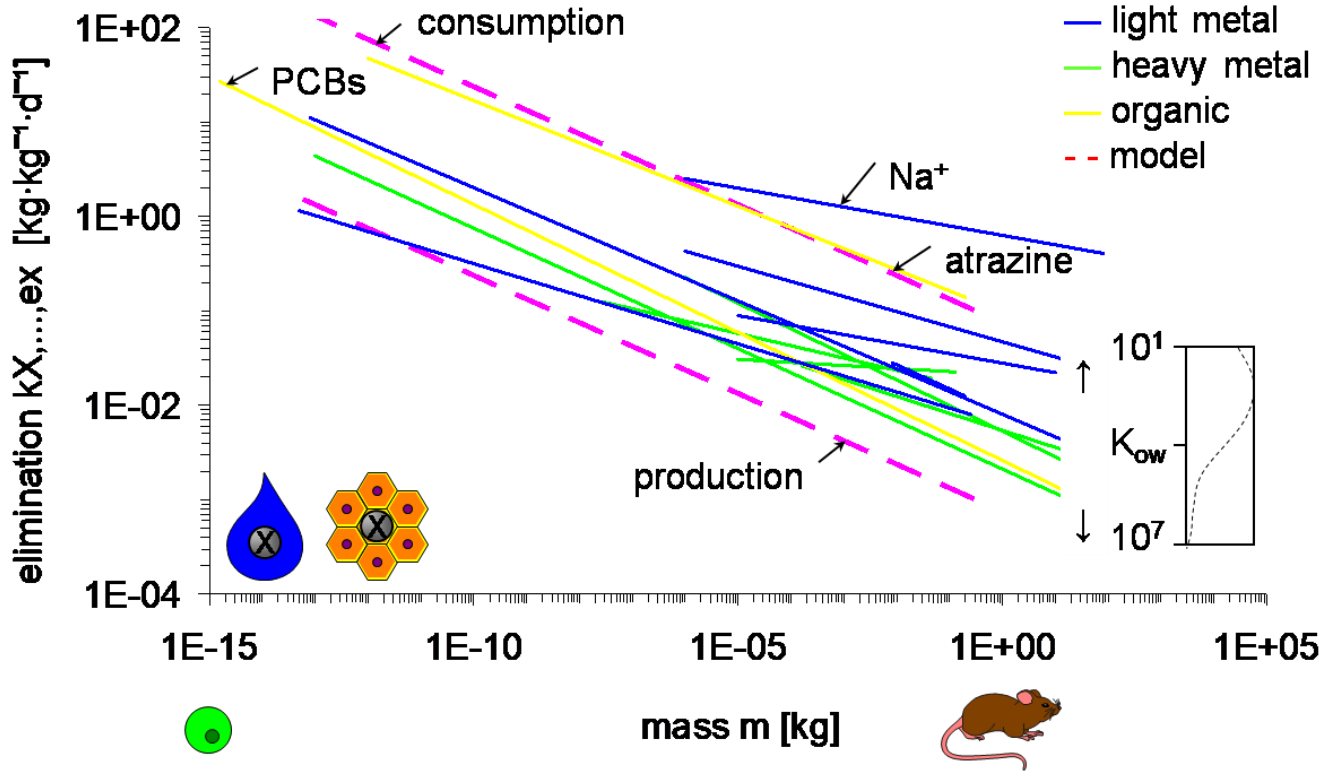


Figure 1. Regressions of elimination rate constants [$kg \cdot kg^{-1} \cdot d^{-1} = d^{-1}$] as a function of organism mass m [kg], ranging from algae to mammals, and organic chemicals' octanol-water partition ratio K_{ow} and metal mass within the limits set by consumption and production (redrawn by author, based on Hendriks et al. 2001).

Scaling of rate constants for uptake and elimination, such as in Figure 1, implies that small organisms reach a given internal concentration faster than large ones. Vice versa, lethal concentrations in water or food needed to reach the same internal level after equal (short-term) exposure duration are lower in smaller compared to larger organisms. Thus, the apparent "sensitivity" of daphnids can, at least partially, be attributed to their small body-size. This emphasizes the need to understand simple scaling relationships before developing to more elaborate explanations.

Using Figure 1, one can, within strict conditions not elaborated here, theoretically relate median lethal concentrations LC_{50} [$\mu g\ L^{-1}$] to the K_{ow} of the chemical and the size of the organism, with $r^2 > 0.8$ (Hendriks, 1995; Table 1). Complicated responses like

susceptibility to toxicants can be predicted only from K_{ow} and body size, which illustrates the generality and power of allometric scaling. Of course, the regressions describe the general trends and in individual cases the deviations can be large. Still, considering the challenges of risk assessment as outlined above, and in the absence of specific data, the predictions in Table 1 can be considered as a reasonable first approximation.

Table 1. Lethal concentrations and doses as a function of test animal body-mass

Species	Endpoint	Unit	b (95% CI)	r^2	n_c	n_s	Source
Guppy	LC_{50}	$mg \cdot L^{-1}$	0.66 (0.51-0.80)	0.98	6	1	1
Mammals	$LD_{10} \approx MTD$	$mg \cdot animal^{-1}$	0.73	0.69-0.77	27	5	2
Birds	Oral LD_{50}	$mg \cdot animal^{-1}$	1.19 (0.67-0.82)	0.76	194	3…37	3
Mammals	Oral LD_{50}	$mg \cdot animal^{-1}$	0.94 (1.18-1.20)	0.89	167	3…16	4
Mammals	Oral LD_{50}	$mg \cdot animal^{-1}$	1.01 (1.00-1.01)		>5000	2…8	5

MTD = maximum threshold dose, repeated dosing. LD_{50} single dose, b = slope of regression line, n_c = number of chemicals, n_s = number of species. Sources: 1 Anderson & Weber (1975), 2 Travis & White (1987), 4 Sample & Arenal (1999), 5 Burzala-Kowalczyk & Jongbloed (2011).

Allometry is also important when dealing with other levels of biological organisation. Leaf or gill area, the number of eggs in ovaries, the number of cell types and many other cellular and organ characteristics scale to body-size as well. Likewise, intrinsic rates of increase (r) of populations and the production-biomass ratios (P/B) of communities can also be obtained from the (average) species mass. Even the area needed by animals in laboratory assays scales to size, i.e., by $m^{¾}$, approximately the same slope noted for home ranges of individuals in the field.

Future perspectives

Since almost any physiological and ecological process in toxicokinetics and toxicodynamics depends on species size, allometric models are gaining interest. Such an approach allows one to quantitatively attribute outliers (like apparently "sensitive" daphnids) to simple biological traits, rather than detailed chemical-toxicological mechanisms.

Scaling has been used in risk assessment at the molecular level for a long time. The molecular size of a compound is often a descriptor in QSARs for accumulation and toxicity. If not immediately evident as molecular mass, volume or area often pops up as an indicator of steric properties. Scaling does not only apply to bioaccumulation and toxicity from molecular to community levels, size dependence is also observed in other sections of the environmental cause-effect chain. Emissions of substances, e.g., scale non-linearly to the size of engines and cities. Concentrations of chemicals in rivers depend on water discharge, which in itself is an allometric function of catchment size. Hence, understanding the principles of cross-disciplinary scaling is likely to pay off in protecting many species against many chemicals.

References

Anderson, P.D., Weber, L.J. (1975). Toxic response as a quantitative function of body size. Toxicology and Applied Pharmacology 33, 471-483.

Burzala-Kowalczyk, L., Jongbloed, G. (2011). Allometric scaling: Analysis of LD50 data. Risk Analysis 31, 523-532.

Hendriks, A.J. (1995). Modelling response of species to microcontaminants: Comparative ecotoxicology by (sub)lethal body burdens as a function of species size and octanol-water partitioning of chemicals. Ecotoxicology and Environmental Safety 32, 103-130.

Hendriks, A.J. (2013). How to deal with 100,000+ substances, sites, and species: Overarching principles in environmental risk assessment. Environmental Science and Technology 47, 3546−3547.

Hendriks, A.J., Van der Linde, A., Cornelissen, G., Sijm, D.T.H.M. (2001). The power of size: 1. Rate constants and equilibrium ratios for accumulation of organic substances. Environmental Toxicology and Chemistry 20, 1399-1420.

Notenboom, J., Vaal, M.A., Hoekstra, J.A. (1995). Using comparative ecotoxicology to develop quantitative species sensitivity relationships (QSSR). Environmental Science and Pollution Research 2, 242-243.

Peters, R.H. (1983). *The Ecological Implications of Body Size*. Cambridge University Press, Cambridge.

Sample, B.E., Arenal, C.A. (1999) Allometric models for interspecies extrapolation of wildlife toxicity data. Bulletin of Environmental Contamination and Toxicology 62, 653-66.

Travis, C.C., White, R.K. (1988) Interspecific scaling of toxicity data. Risk Analysis 8, 119-125.

4.1.5. Question 1

Think of a biological trait that you are interested in. How do you think it will scale to organism size?

4.1.5. Question 2

What is the elimination rate constant of a chemical with a K_{ow} of 10^4 from an insect with a mass of 10^{-5} kg?

4.1.5. Question 3

Why is size-scaling more prominent in toxicokinetics than in toxicodynamics?

4.1.6. Food chain transfer

Auteur: Nico van den Brink

Reviewers: Kees van Gestel. Jan Hendriks

Learning objectives:

You should be able to:

- Mention the chemical properties determining the potential of chemicals to accumulate in food chains
- Explain the role of biological and ecological factors in the food-chain accumulation of chemicals

Keywords: biomagnification, food-chain transfer,

Accumulation of chemicals across different trophic levels.

Chemicals may be transferred from one organism to another. Grazers will ingest chemicals that are in the vegetation they eat. Similarly, predators are exposed to chemicals in their prey items. This so-called food web accumulation is governed by properties of the chemical, but also by some traits of the receiving organism (e.g. grazer or predator).

Chemical properties driving food web accumulation

Some chemicals are known to accumulate in food webs, reaching the highest concentrations in top-predators. Examples of such chemicals are organochlorine pesticides like DDT and brominated flame retardants (e.g. PBDEs; see section on POPs). Such accumulating chemicals have a few properties in common: they need to be persistent and they need to have affinity for the organismal body. Organic chemicals with a relatively high log Kow, indicating a high affinity for lipids, will enter organisms quite effectively (see section on Bioconcentration and kinetics modelling). Once in the body, these chemicals will be distributed to lipid rich tissues, and excretion is rather limited. In case of persistent chemicals that are not metabolised, concentrations will increase over time when uptake is higher than excretion. Furthermore, such chemicals are likely to be passed on to organisms at the next trophic level in case of prey-predator interactions. Some of these chemicals may be metabolised by the organism, most often into more water soluble metabolites (see section on Xenobiotic metabolism & defence). These metabolites are more easily excreted, and in this way concentrations of metabolizable chemicals do not increase so much over time, and will therefore also transfer less to higher trophic levels. The effects of metabolism on the internal concentrations of organisms is clearly illustrated by a study on the

uptake of organic chemicals by different aquatic species (Kwok et al., 2013). In that study the uptake of persistent chemicals (organochlorine pesticides; OCPs) was compared with the uptake of chemicals that may be metabolised (polycyclic aromatic hydrocarbons; PAHs). The authors compared shrimps with fish, the former having a limited capacity to metabolise PAHs while fish can. Figure 1 shows the Biota-to-Sediment Accumulation Factors (BSAFs; see section on Bioaccumulation), which is the ratio between the concentration in the organism and in the sediment. It is shown that OCPs accumulate to a high extent in both species, reflecting persistent, non-metabolizable chemicals. For PAHs the results are different per species, fish are able to metabolise and as a result the concentrations of PAHs in the fish are low, while in shrimp, with a limited metabolic capacity, the accumulation of PAHs is comparable to the OCPs. These results show that not only the properties of the chemicals are of importance, but also some traits of the organisms involved, in this case the metabolic capacity.

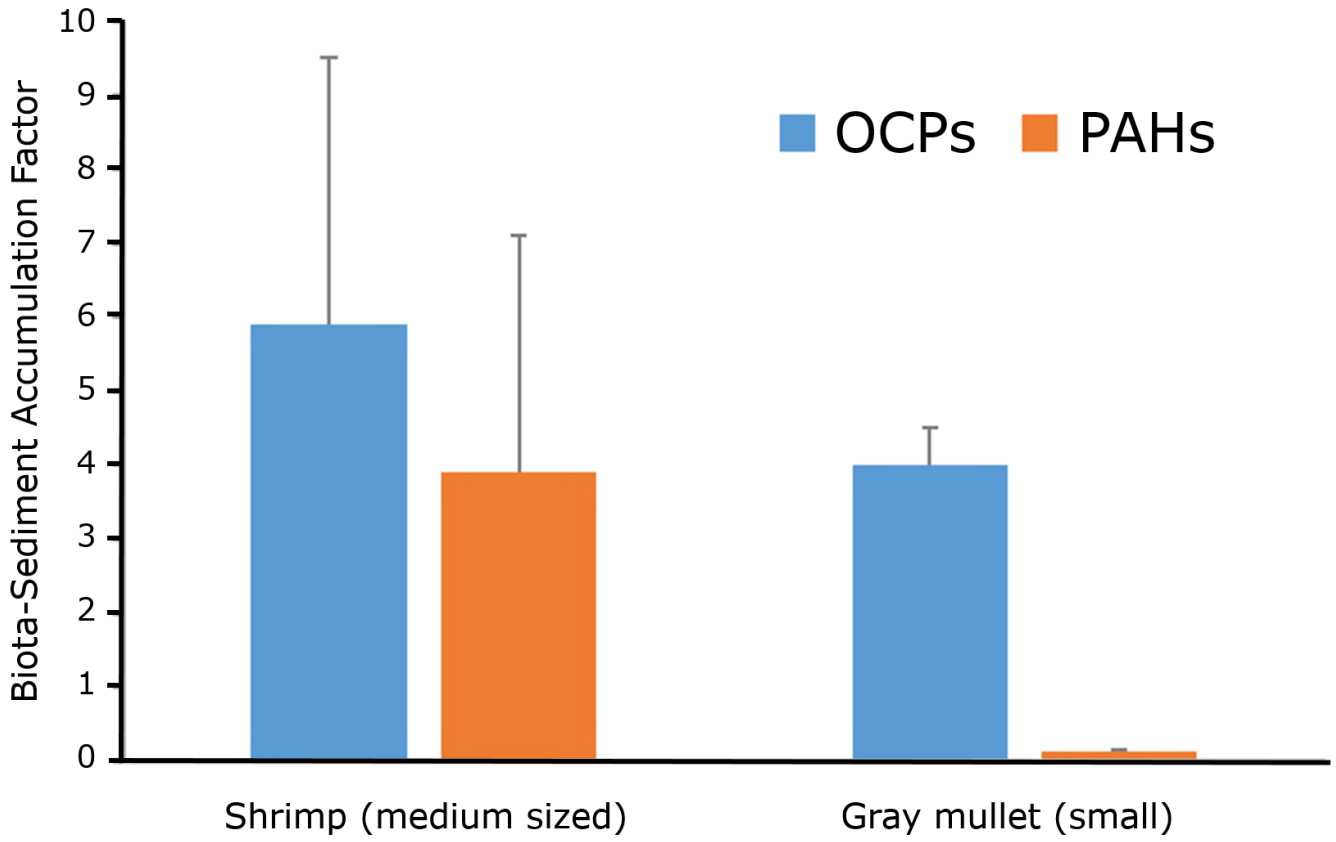


Figure 1. Biota-sediment accumulation factors organochlorine pesticides (OCPs) and polycyclic aromatic hydrocarbons (PAHs) in fish and shrimp. Redrawn from Kwok et al. (2013).

Effects of species traits and food web structure on food web accumulation

Food-web accumulation of chemicals is driven by food uptake. At lower trophic levels, most organisms will acquire relatively low concentrations from the ambient environment. First consumers, foraging on these organisms will accumulate the chemicals of all of them, and in case of persistent chemicals that enter the body easily, concentrations in the consumers will be higher than in their diet. Similarly, concentrations will increase when the chemicals are transferred to the next trophic level. This process is called biomagnification, indicating increasing concentrations of persistent and accumulative chemicals in food webs. The most iconic example of this is on the increasing concentrations of DDTs in fish eating American Osprey (Figure 2), a casus which has led to the ban of a lot of organochlorine chemicals.

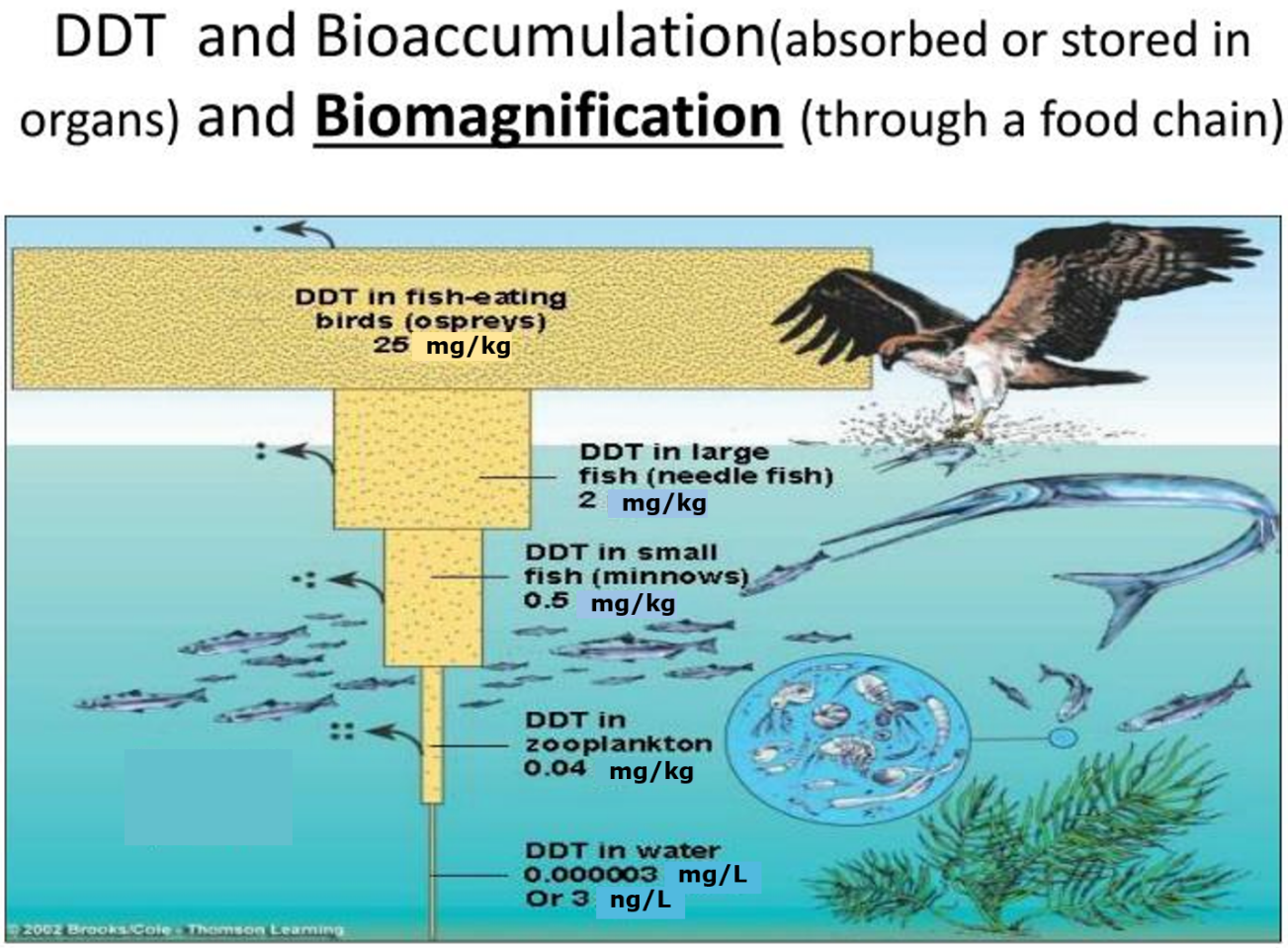


Figure 2. Example of biomagnification of DDT in an aquatic food web of ospreys. Slightly modified from: http://naturalresources.anthro-semin...magnification/.

Since biomagnification along trophic levels is food driven, it is of importance to include diet composition into the studies. This can be explained by an example on small mammals in the Netherlands. Two similar small mammal species, the bank vole (*Myodes glareolus*) and the common vole (*Microtus arvalis*) co-occur in larger part of the Netherlands. Although the species look very similar, they are different in their diet and habitat use. The bank vole is a omnivorous species, inhabiting different types of habitat while the common vole is strictly vegetarian living in pastures. In a study on the species-specific uptake of cadmium, diet items of both species were analysed, indicating nearly 3 orders of magnitudes differences in cadmium concentrations between earthworms and berries from vegetation (Fig 3A, van den Brink et al., 2010). Stable isotopic ratios of carbon and nitrogen were used to assess the general diets of the organisms. The common vole ate mostly stinging nettle and grass, including seeds, while the bank vole showed to forage on grass herbs and earthworms. This difference in diet was reflected in increased concentrations of cadmium in the bank vole in comparison to the common vole (both inhabiting the same area). The concentrations of one bank vole appeared to be extremely low (red diamond in Figure 3b), and initially this was considered to be an artefact. However, detailed analysis of the stable isotopic ratios in this individual revealed that it had foraged on stinging nettle and grass, hence a diet more reflecting the common vole. This emphasises once more that organisms accumulate through their diet (you accumulate what you eat!)

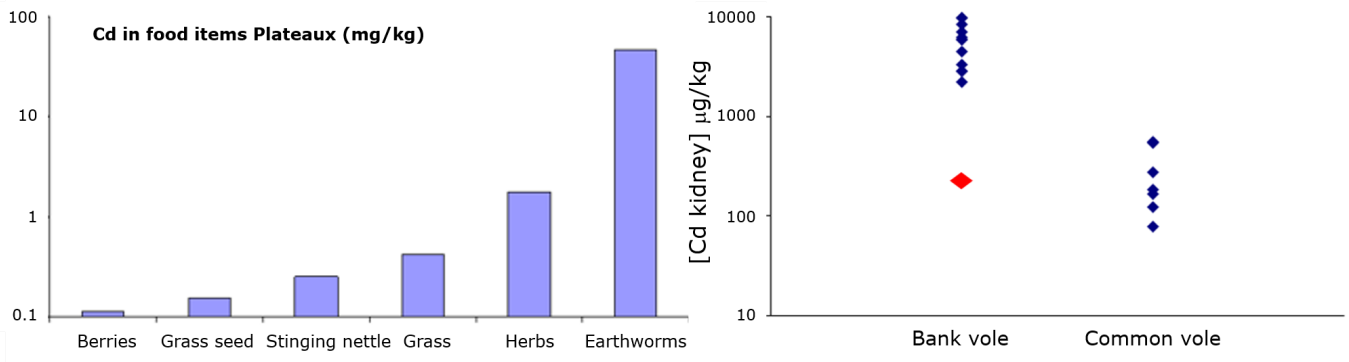


Figure 3. Left: Concentrations of cadmium in diet items of small mammals collected in the Plateaux area near Eindhoven, the Netherlands; Right: Cadmium concentrations in kidneys of bank voles and common voles collected in the Plateaux area. See text for further explanation. Redrawn from van den Brink et al. (2010).

Case studies

Orcas or Killer whales (*Orcinus orca*) are large marine predatory mammals, which roam all around the oceans, from the Arctic to the deep south region of the Antarctic. Although they appear ubiquitous around the world, generally different pods of Orcas occur in different regions of the marine ecosystem. Often, each pod has developed specialised foraging behaviours targeted at specific prey species. Although Orcas are generally apex top-predators at the top of the (local) food web, the different foraging strategies would suggest that exposure to accumulating chemicals may differ considerably between pods. This was indeed shown to be the case in a very elaborate study on different pods of Orcas of the West-coast of Canada by Ross et al. (2000). In the Vancouver region there is a resident pod while the region is also often visited by two transient groups of Orcas. PCB concentrations were high in all animals, but the transient animals contained significantly higher levels. The transient whales mainly fed on marine mammals, while the resident animals mainly fed on fish, and this difference in diet was thought to be the cause of the differences in PCB levels between the groups. In that study, it was also shown that PCB levels increased with age, due to the persistence of the PCBs, while female orcas contained significant lower concentrations of PCBs. The latter is caused by the lactation of the female Orcas during which they feed their calves with lipid rich milk, containing relatively high levels of (lipophilic) PCBs. By this process, females offload large parts of their PCB body burden, however by transferring these PCBs to their developing calves (see also Figure 3 in the section on Bioaccumulation). A recent study showed that although PCBs have been banned for decades now, they still pose threats to populations of Orcas (Deforges et al., 2018). In that study, regional differences in PCB burdens were confirmed, likely due to differences in diet preferences although not specifically mentioned. It was shown that PCB levels in most of the Orca populations were still above toxic threshold levels and concerns were raised regarding the viability of these populations. This study confirms that 1) Orcas are exposed to different levels of PCBs according to their diet, which influences the biomagnification of the PCBs, 2) Orca populations are very inefficient in clearing PCBs from the individual due to little metabolism but also from the population due to the efficient maternal transfer from mother to calve, and 3) persistent, accumulating chemicals may pose threats to organisms even decades after their use. Understanding the mechanisms and processes underlying the biomagnification of persistent and toxic compounds is essential for a in depth risk assessment.

References

Desforges, J.-P., Hall, A., McConnell, B., Rosing-Asvid, A., Barber, J.L., Brownlow, A., De Guise, S., Eulaers, I., Jepson, P.D., Letcher, R.J., Levin, M., Ross, P.S., Samarra, F., Víkingson, G., Sonne, C., Dietz, R. (2018). Predicting global killer whale population collapse from PCB pollution. Science 361, 1373-1376.

Ford, J.K.B., Ellis, G.A., Matkin, D.R., Balcomb, K.C., Briggs, D., Morton, A.B. (2005). Killer whale attacks on minke whales: Prey capture and antipredator tactics. Marine Mammal Science 21, 603-618.

Guinet, C. (1992). Predation behaviour of killer whales (*Orcinus orca*) around Grozet Islands. Canadian Journal of Zoology 70, 1656-1667.

Kwok, C.K., Liang, Y., Leung, S.Y., Wang, H., Dong, Y.H., Young, L., Giesy, J.P., Wong, M.H. (2013). Biota-sediment accumulation factor (BSAF), bioaccumulation factor (BAF), and contaminant levels in prey fish to indicate the extent of PAHs and

OCPs contamination in eggs of waterbirds. Environmental Science and Pollution Research 20, 8425-8434.

Ross, P.S., Ellis, G.M., Ikonomou, M.G., Barrett-Lennard, L.G., Addison, R.F. (2000). High PCB concentrations in free-ranging Pacific killer whales, *Orcinus orca*: Effects of age, sex and dietary preference. Marine Pollution Bulletin 40, 504-515.

Samarra, F.I.P., Bassoi, M., Beesau, J., Eliasdottir, M.O., Gunnarsson, K., Mrusczok, M.T., Rasmussen, M., Rempel, J.N., Thorvaldsson, B., Vikingsson, G.A. (2018). Prey of killer whales (*Orcinus orca*) in Iceland. Plos One 13, 20.

van den Brink, N., Lammertsma, D., Dimmers, W., Boerwinkel, M.-C., van der Hout, A. (2010). Effects of soil properties on food web accumulation of heavy metals to the wood mouse (*Apodemus sylvaticus*). Environmental Pollution 158, 245-251.

4.1.6. Question 1

In a lake, the ecosystem consists of algae, daphnids feeding on the algae and fish feeding on the daphnids. Chemical A has a low log K_{ow}, but is persistent. Chemical B has high log K_{ow}, but can be metabolised by daphnids and fish, and Chemical C has high log K_{ow} and is also persistent. Where may you find the highest concentrations of Chemical A, B and C?

4.1.6. Question 2

Name the two traits of species determining the potential for chemicals to reach high concentrations in it.

4.1.6. Question 3

This is a more general question, trying to put the information in the text of this section on bioaccumulation across different trophic levels in a broader perspective:

What property does a chemical need, besides high potential for bioaccumulation and persistence, to have a high likelihood to pose serious environmental risks to organisms?

4.1.7. Critical Body Concentration

Author: Martina G. Vijver

Reviewers: Kees van Gestel and Frank Gobas

Learning objectives:

You should be able to

- describe the Critical Body Concentration (CBC) concept for assessing the toxicity of chemicals.
- graphically explain the CBC concept and make a distinction between slow and fast kinetics.
- mention cases in which the CBC approach fails.

Keywords:

Time dependent effects, internal body concentrations, one compartment model

Introduction

One of the quests in ecotoxicology is how to link toxicity to exposure and to understand why some organisms experience toxic effects while others do not at the same level of exposure. A generally accepted approach for assessing possible adverse effects on biota, no matter what kind of species, is the Critical Body Concentration (CBC) concept (McCarty 1991). According to this concept, toxicity is determined by the amount of chemical taken up, so by the internal concentration, which has a relationship with the duration of the exposure as well as the exposure concentration.

Figure 1 shows the relationship between the development with time of the internal concentrations of a chemical in an organism and the time when mortality occurs at different exposure concentrations under constant exposure. Independent of exposure time or exposure concentration, mortality occurs at a more or less fixed internal concentration. The CBC is defined as the highest internal concentration of a substance in an organism that does cause a defined effect, e.g. 50% mortality or 50% reduction in the number of offspring produced. By comparing internal concentrations measured in exposed organisms to CBC values derived in the laboratory, a measure of risk is obtained. The CBC applies to lethality as well as to sub-lethal effects like reproduction or growth inhibition.

Relating toxicity to toxicokinetics

From Figure 1A, it may also become clear that chemicals that have fast uptake kinetics will reach the CBC faster than chemicals that have slow kinetics (see Section on Bioaccumulation kinetics). As a consequence, also the time to reach a constant LC_{50} (indicated as the ultimate LC_{50}: LC_{50}¥; Figure 1B) depends on kinetics. Hence, both toxic effects and chemical concentration are controlled by the same kinetics.The CBC can be derived from the LC_{50}-time relationship and be linked to the LC_{50}¥ using uptake and elimination rate constants (k_1 and k_2). It should be noted that the k_2 in this case does not reflect the rate of chemical excretion but rather the rate of elimination of toxic effects caused by the chemical (so, note the difference here with Section on Bioaccumulation kinetics).

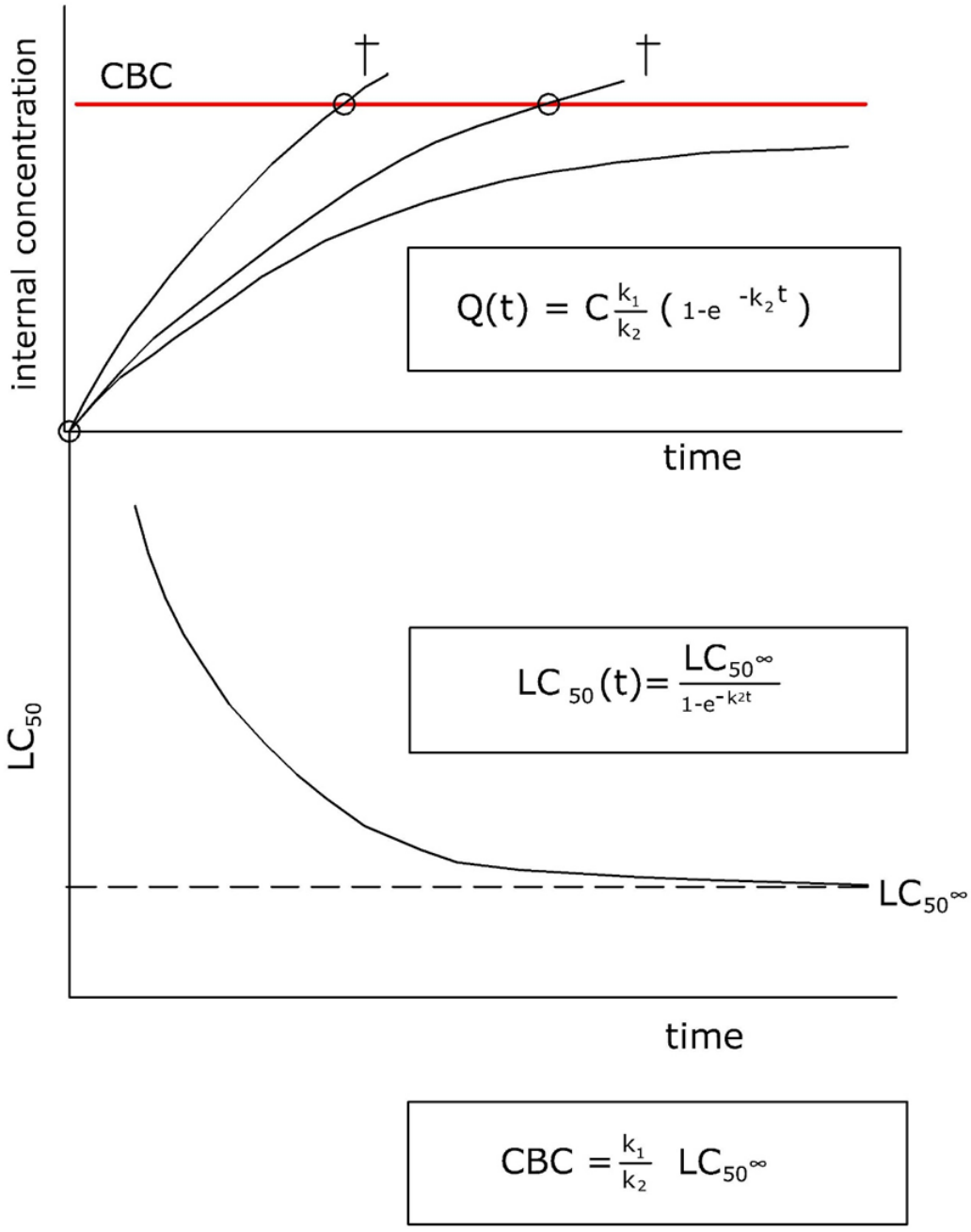




Figure 1: A. The relationship between uptake kinetics of a chemical in an organism and its toxicity according to the Critical Body Concentration (CBC) concept under constant exposure. The red line depicts the highest internal concentration of the chemical in the organism killing the organism (CBC); exceedance of that line results in mortality. B. The relationship between LC_{50} and time. LC_{50} will reach a constant value with time, indicated as the ultimate LC_{50} (LC_{50}¥). The CBC can be calculated from the LC_{50}¥ using uptake and elimination rate constants (k_1 and k_2) derived by first order kinetics as shown here. Note: the CBC approach is not limited to first order kinetics. And the same curves may be seen when focusing on sublethal effects, LC_{50} then reads EC_{50}. Drawn by Wilma Ijzerman.

The time needed to reach steady state depends on the body size of the organisms, with larger organisms taking longer time to attain steady state compared to smaller organisms (McCarty 1991). The time needed to reach steady state depends also on the exposure surface area of the exposed organisms (Pawlisz and Peters 1993) as well as their metabolic activity. Organisms not capable of excreting or metabolizing a chemical will continue accumulating with time, and the LC_{50}¥ will be zero. This is e.g. the case for cadmium in isopods (Crommentuijn et al. 1994), but kinetics are so slow that cadmium in these animals never reaches lethal concentrations in relatively clean environments as their life span is too short.

The CBC integrates environmentally available fractions with bioavailable concentrations and toxicity at specific receptors (McCarty and MacKay 1993). See also Section on Bioavailability. In this way, the actual exposure concentration in the environment does not need to be known for performing a risk assessment. The internal concentration of the chemical in the organism is the only concentration required for a risk assessment. Therefore many difficulties are overcome regarding bioavailability issues, e.g. it removes some of the disadvantages of the exposure concentration expressed per unit of soil, as well as of dealing with exposures that vary over time or space.

Proof of the CBC concept

A convincing body of evidence was collected to support the CBC approach. For organic compounds with a narcotic mode of action, effects could be assessed over a wide range of organisms, test compounds and exposure media. For narcotic compounds with octanol-water partition coefficients (K_{ow}) varying from 10 to 1,000,000 (see for details Section on Relevant chemical properties), the concentration of chemical required for lethality through narcosis is approximately 1-10 mmol/kg: Figure 2 (McCarty and MacKay 1993).

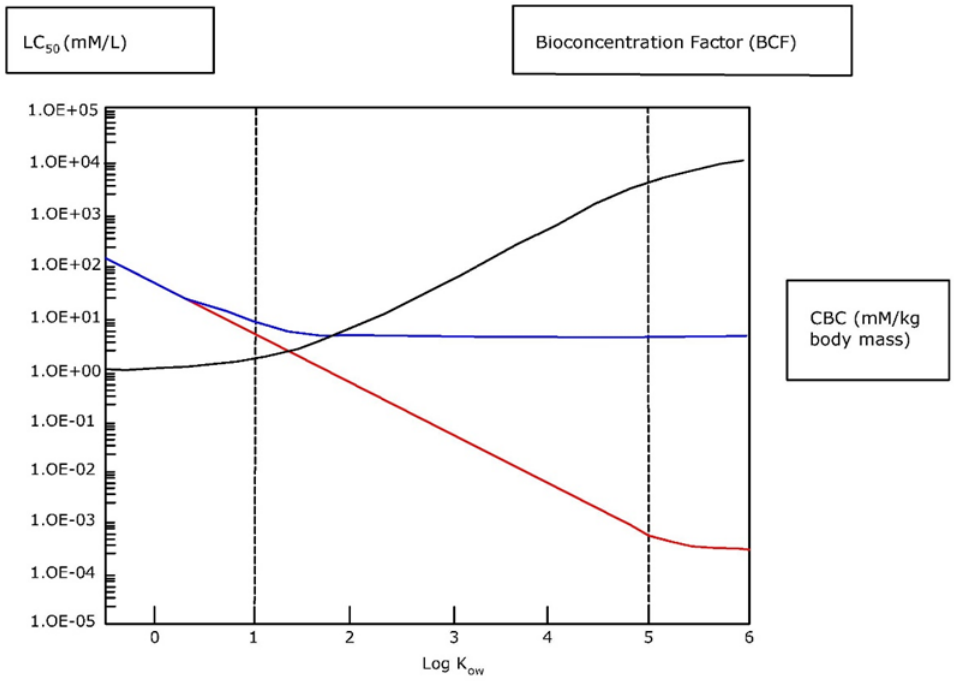


Figure 2: Theoretical plot supporting the Critical Body Concentration (CBC) concept for non-polar organic chemicals acting by narcosis. The bioconcentration factor (BCF in L/kg body mass; black line) of these chemicals increases with increasing log K_{ow} while their acute toxicity also increases (LC_{50} in test solution in mM decreases; red line). The product of LC_{50} and BCF is the critical body concentrations (in mM/kg body mass; blue line). Adapted from McCarty and Mackay (1993) by Wilma Ijzerman.

To reduce the variation in bioconcentration factor (BCF) values for the accumulation of chemicals in organisms from water, normalization by lipid content has been suggested allowing to determine the chemical activity within an organism's body (US EPA 2003). For that reason, lipid extraction protocols are intensively described within the updated OECD Guideline for the testing of chemicals No. 305 for fish bioaccumulation tests, along with a sampling schedule of lipid measurement in fish. Correction of the BCF for differences in lipid content is also described in the same OECD guideline No. 305. If chemical and lipid analyses have been conducted on the same fish, this requires correction for the corresponding lipid content of each individual measured concentration in the fish. This should be done *prior* to using the data to calculate the kinetic BCF. If lipid content is not measure on all sampled fish, a mean lipid content of approx. 5% must be used to normalize the BCF. It should be noted that this correction holds only for chemicals accumulating in lipids and not for chemicals that do primarily bind to proteins (e.g. perfluorinated substances).

When does the CBC concept not apply?

The CBC concept also has some limitations. Crommentuijn et al. (1994) found that the toxicity of metals to soil invertebrates could not be explained using critical body concentrations. The way different organisms deal with accumulated metals has a large impact on the magnitude of body concentrations reached and the accompanying metal sensitivity (Rainbow 2002). Moreover, adaptation or development of metal tolerance limits the application of CBCs for metals. When the internal metal concentration does not show a monotonic relationship with the exposure concentration, it is not possible to derive CBCs. This means that whenever organisms are capable of trapping a portion of the metal in forms that are not biologically reactive, a direct relationship between body metal concentrations and toxicity may be absent or less evident (Luoma and Rainbow 2005, Vijver et al. 2004). Consequently, for metals a wide range of body concentrations with different biological significance exists. It therefore remains an open question whether the

approach is applicable to modes of toxic action other than narcosis. Another important point is the question to what extent the CBC approach is applicable to assessing the effect of chemical mixtures, especially in cases the chemicals have a different mode of action.

References

Crommentuijn, T., Doodeman, C.J.A.M., Doornekamp, A., Van der Pol, J.J.C., Bedaux, J.J.M., Van Gestel, C.A.M. (1994). Lethal body concentrations and accumulation patterns determine time-dependent toxicity of cadmium in soil arthropods. Environmental Toxicology and Chemistry 13, 1781-1789.

Luoma, S.N., Rainbow, P.S. (2005). Why is metal bioaccumulation so variable? Biodynamics as a unifying concept. Environmental Science and Technology 39, 1921-1931

McCarty, L.S. (1991). Toxicant body residues: implications for aquatic bioassays with some organic chemicals. In: Mayes, M.A., Barron, M.G. (Eds.), Aquatic Toxicology and Risk Assessment: Fourteenth Volume. ASTM STP 1124. Philadelphia: American Society for Testing and Materials. pp. 183-192. DOI: 10.1520/STP23572S

McCarty, L.S., Mackay, D. (1993). Enhancing ecotoxicological modeling and assessment. Environmental Science and Technology 27, 1719-1727

Pawlisz, A.V., Peters, R.H. (1993). A test of the equipotency of internal burdens of nine narcotic chemicals using *Daphnia magna*. Environmental Science and Technology 27, 2801-2806

Rainbow P.S. (2002). Trace metal concentrations in aquatic invertebrates: why and so what? Environmental Pollution 120, 497-507.

U.S. EPA. (2003). In: Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health: Technical Support Document. Volume 2. United States Environmental Protection Agency, Washington, D.C: Development of National Bioaccumulation Factors.

Vijver M.G., Van Gestel, C.A.M., Lanno, R.P., Van Straalen, N.M., Peijnenburg, W.J.G.M. (2004) Internal metal sequestration and its ecotoxicological relevance: a review. Environmental Science and Technology 18, 4705-4712.

4.1.7. Question 1

Explain why the CBC approach integrates chemical and biological availability.

4.1.7. Question 2

How are time dynamics involved in the CBC approach?

4.1.7. Question 3

Under what conditions the CBC approach cannot be applied?

This page titled 2.3.4.1: Toxicokinetics is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.1: Toxicokinetics** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.4.2: Toxicodynamics and Molecular Interactions

4.2. Toxicodynamics & Molecular Interactions

Author: Timo Hamers

Reviewers: Frank van Belleghem and Ludek Blaha

Learning goals

You should be able to

- explain that a toxic response requires a molecular interaction between a toxic compound and its target
- name at least three different types of biomolecular targets
- name at least three functions of proteins that can be hampered by toxic compounds
- explain in general terms the consequences of molecular interaction with a receptor protein, an enzyme, a transporter protein, a DNA molecule, and a membrane lipid bilayer.

Key words: Receptor; Transcription factor; DNA adducts; Membrane; Oxidative stress

Description

Toxicodynamics describes the dynamic interactions between a compound and its biological target, leading ultimately to an (adverse) effect. In this Chapter 4.2, toxicodynamics have been described for processes leading to diverse adverse effects. Any adverse effects by a toxic substance is the result of an interaction between the toxicant and its biomolecular target (i.e. mechanism of action). Biomolecular targets include a protein, a DNA or RNA molecule, a phospholipid bilayer membrane, but also small molecules that have specific functions in keeping cellular homeostasis.

Both endogenous and xenobiotic compounds that bind to proteins are called ligands. The consequence of a protein interaction depends on the role of the target protein, e.g.

1. Receptor

2. Enzyme

3. Protein

Receptor proteins specifically bind and respond to endogenous signalling ligands such as hormones, prostaglandins, growth factors, or neurotransmitters, by causing a typical cellular response. Receptor proteins can be located in the cell membrane, in the cytosol, and in the nucleus of a cell. Agonistic receptor ligands activate the receptor protein whereas antagonistic ligands inactivate the receptor and prevent (endogenous) agonists from activating the receptor. Based on the role of the receptor protein, binding by ligands may interfere with ion channels, G-protein coupled receptors, enzyme linked receptors, or nuclear receptors. Xenobiotic ligands can interfere with these cellular responses by acting as agonistic or antagonistic ligands (link to section on Receptor interaction).

Compounds that bind to an enzyme usually cause inhibition of the enzyme activity, i.e. a decrease in the conversion rate of the endogenous substrate(s) of the enzyme into its/their corresponding product(s). Compounds that bind non-covalently to an enzyme cause reversible inhibition, while compounds that bind covalently to an enzyme cause irreversible inhibition (link to section on Protein inactivation).

Similarly, compounds that bind to a transporter protein usually inhibit the transport of the natural, endogenous ligand. Such transporter proteins may be responsible for local transport of endogenous ligands across the cell membrane, but also for peripheral transport of endogenous ligands through the blood from one organ to the other (link to section Endocrine disruption).

Apart from interaction with functional receptor, enzyme, or transporter proteins, toxic compounds may also interact with structural proteins. For instance the cytoskeleton may be damaged by toxic compounds that block the polymerization of actin, thereby preventing the formation of filaments.

In addition to proteins, DNA and RNA macromolecules can be targets for compound binding. Especially the guanine base can be covalently bound by electrophilic compounds, such as reactive metabolites. Such DNA adducts may cause copy errors during DNA

replication leading to point mutations (link to section on Genotoxicity).

Compounds may also interfere with phospholipid bilayer membranes, especially with the outer cell membrane and with mitochondrial membranes. Compounds disturb the membrane integrity and functioning by partitioning into the lipid bilayer. Lost membrane integrity may ultimately lead to leakage of electrolytes and loss of membrane potential.

Narcosis and Membrane Damage
Partitioning into the lipid bilayer is a non-specific process. Therefore, concentrations in biological membranes that cause effects through this mode of action do not differ between compounds. As such, this type of toxicity is considered as a "baseline toxicity" (also called "narcosis"), which is exerted by all chemicals. For instance, the chemical concentration in a target membrane causing 50% mortality in a test population is around 50 mmol/kg lipid, irrespective of the species or compound under consideration. Based on external exposure levels, however, compounds do have different narcotic potencies. After all, to reach similar lipid-based internal concentrations, different exposure concentrations are required, depending on the lipid-water partitioning coefficient, which is an intrinsic property of a compound, and not of the species.
Narcotic action is not the only mechanism by which compounds may damage membrane integrity. Compounds called "ionophores", for instance, act like ion carriers that transport ions across the membrane, thereby disrupting the electrolyte gradient across the membrane. Ionophores should not be confused with compounds that open or close ion channels, although both type of compounds may disrupt the electrolyte gradient across the membrane. The difference is that ionophores dissolve in the bilayer membrane and shuttle transport ions across the membrane themselves, whereas ion channel inhibitors or stimulators close or open, respectively, a protein channel in the membrane that acts as a gate for ion transport.

Finally, it should be mentioned here that some compounds may cause oxidative stress by increasing the formation of reactive oxygen species (ROS), such as H_2O_2, O_3, $O_2^{\bullet-}$, •OH, NO•, or RO•. ROS are oxygen metabolites that are found in any aerobic living organism. Compounds may directly cause an increase in ROS formation by undergoing redox cycling or interfering with the electron transport chain. Alternatively, compounds may cause an indirect increase in ROS formation by interference with ROS-scavenging antioxidants, ranging from small molecules (e.g. glutathione) to proteins (e.g. catalase or superoxide dismutase). For compounds causing both direct or indirect oxidative stress, it is not the compound itself that has a molecular interaction with the target, but the ROS which may bind covalently to DNA, proteins, and lipids (link to section on Oxidative Stress).

4.2. Question 1

Name three biomolecular targets that can be affected by a compound

---

4.2. Question 2

Name three different mechanisms by which a compound can affect analyte transport across the cell membrane

---

4.2. Question 3

What is the difference between a receptor agonist and a receptor antagonist?

4.2.1. Protein Inactivation

Author: Timo Hamers

Reviewers: Frank van Belleghem and Ludek Blaha

Learning objectives:

You should be able to

- discuss how a compound that binds to a protein may inhibit ligand binding, and thereby hamper the function of the protein
- explain the mechanism of action of organophosphate insecticides inhibiting acetylcholinesterase
- explain the mechanism of action of halogenated phenols inhibiting thyroid hormone transport by transthyretin
- distinguish between reversible and irreversible protein inactivation
- distinguish between competitive, non-competitive, and uncompetitive enzyme inhibition

Key words: enzyme inhibition; acetylcholinesterase, transthyretin, competitive inhibition, non-competitive inhibition, uncompetitive inhibition

Introduction

Proteins play an important role in essential biochemical processes including catalysis of metabolic reactions, DNA replication and repair, transport of messengers (e.g. hormones), or receptor responses to such messengers. Many toxic compounds exert their toxic action by binding to a protein and thereby disturbing these vital protein functions.

Inhibition of the protein transport function

Binding of xenobiotic compounds to a transporter protein may hamper binding of the natural ligand of the protein, thereby inhibiting the transporter function of the protein. An example of such inhibition is the binding of halogenated phenols to transthyretin (TTR). TTR is a transport protein for thyroid hormones, present in the blood. It has two binding places for the transport of thyroid hormone, i.e. mainly thyroxine (T4) in mammals and mainly triiodothyronine (T3) in other vertebrates (Figure 1). Compounds with high structural resemblance with thyroid hormone (especially halogenated phenols, such as hydroxylated metabolites of PCBs or PBDEs), are capable to compete with thyroid hormone for TTR binding. Apart from the fact that this enhances distribution of the toxic compounds, this also causes an increase of unbound thyroid hormone in the blood, which is then freely available for uptake in the liver, metabolic conjugation, and urinary excretion. Ultimately, this may lead to decreased thyroid hormone levels in the blood.

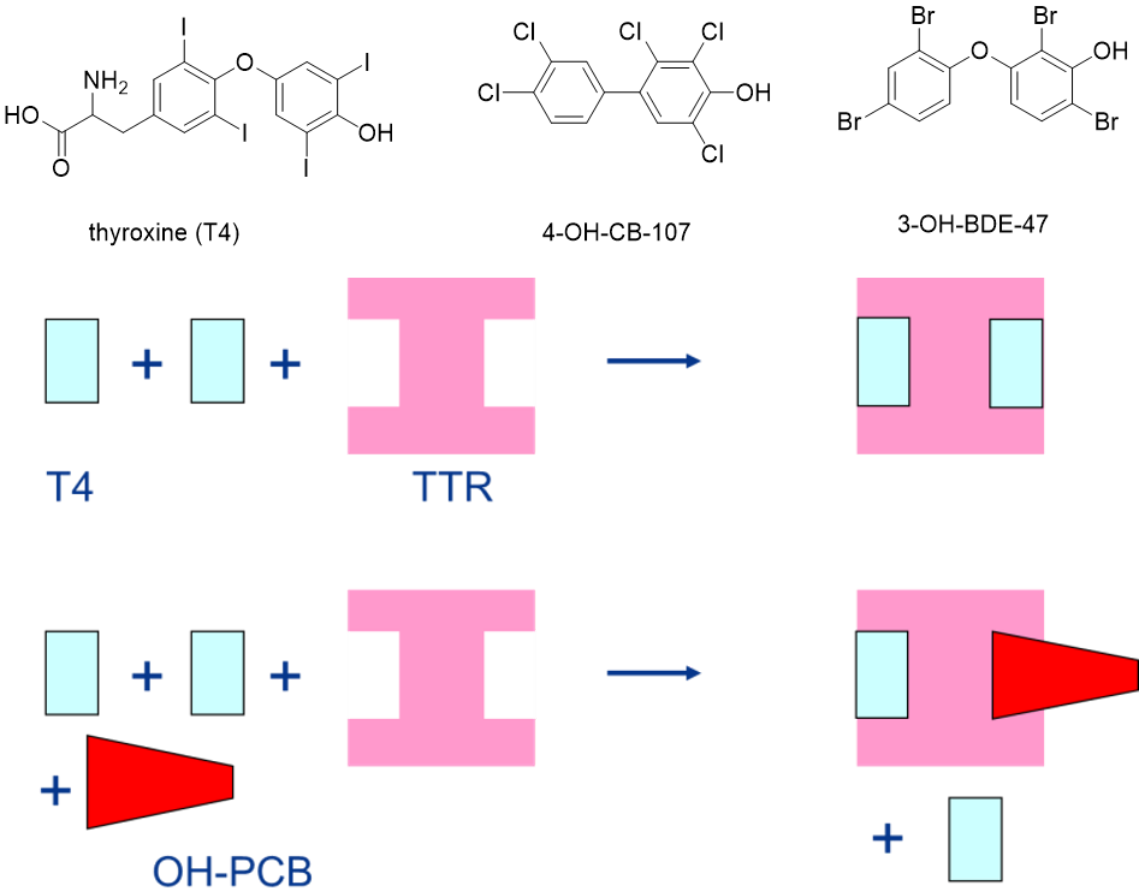


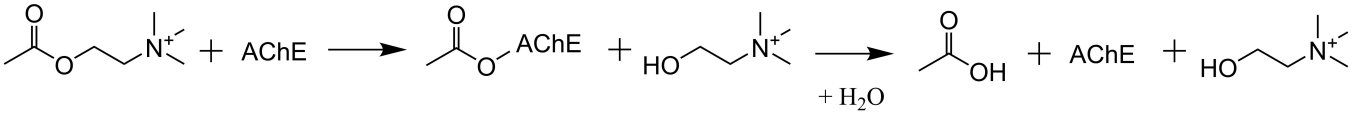
Figure 1: Structural resemblance between T4, a hydroxylated PCB metabolite (4-OH-CB-107) and a hydroxylated PBDE metabolite (3-OH-BDE-47). The lower panel illustrates how halogenated phenols (red; e.g. OH-PCB), given their structural resemblance with T4, can compete with T4 (cyan) for TTR-binding (pink), thereby increasing the levels of unbound T4.

Inhibition of the protein enzymatic activity

Proteins involved in the catalysis of a metabolic reaction are called enzymes. The general formula of such a reaction is

$$Substrate \xrightarrow{enzyme} Product$$

Binding of a toxic compound to an enzyme usually causes an inhibition of the enzyme activity, i.e. a decrease in the conversion rate of the endogenous substrate(s) of the enzyme into its/their corresponding product(s). In practice, this causes a toxic response due to a surplus of substrate and/or a deficit of product. One of the classical examples of enzyme inhibition by toxic compounds is the inhibition of the enzyme acetylcholinesterase (AChE) by organophosphate insecticides. AChE catalyzes the hydrolysis of the neurotransmitter acetylcholine (ACh), in the cholinergic synapses. During transfer of an action potential from one cell to the other, ACh is released in these synapses from the presynaptic cell into the synaptic cleft in order to stimulate the acetylcholine-receptor (AChR) on the membrane of the postsynaptic cell. AChE, which is also present in these synapses, is then responsible to break down the ACh into acetic acid and choline:



By covalent binding to serine residues in the active site of the AChE enzyme, organophosphate insecticides can inhibit this reaction causing accumulation of the ACh neurotransmitter in the synapse (Fig. 2). As a consequence, the AChR is overstimulated causing convulsions, hypertension, muscle weakness, salivation, lacrimation, gastrointestinal problems, and slow heartbeat.

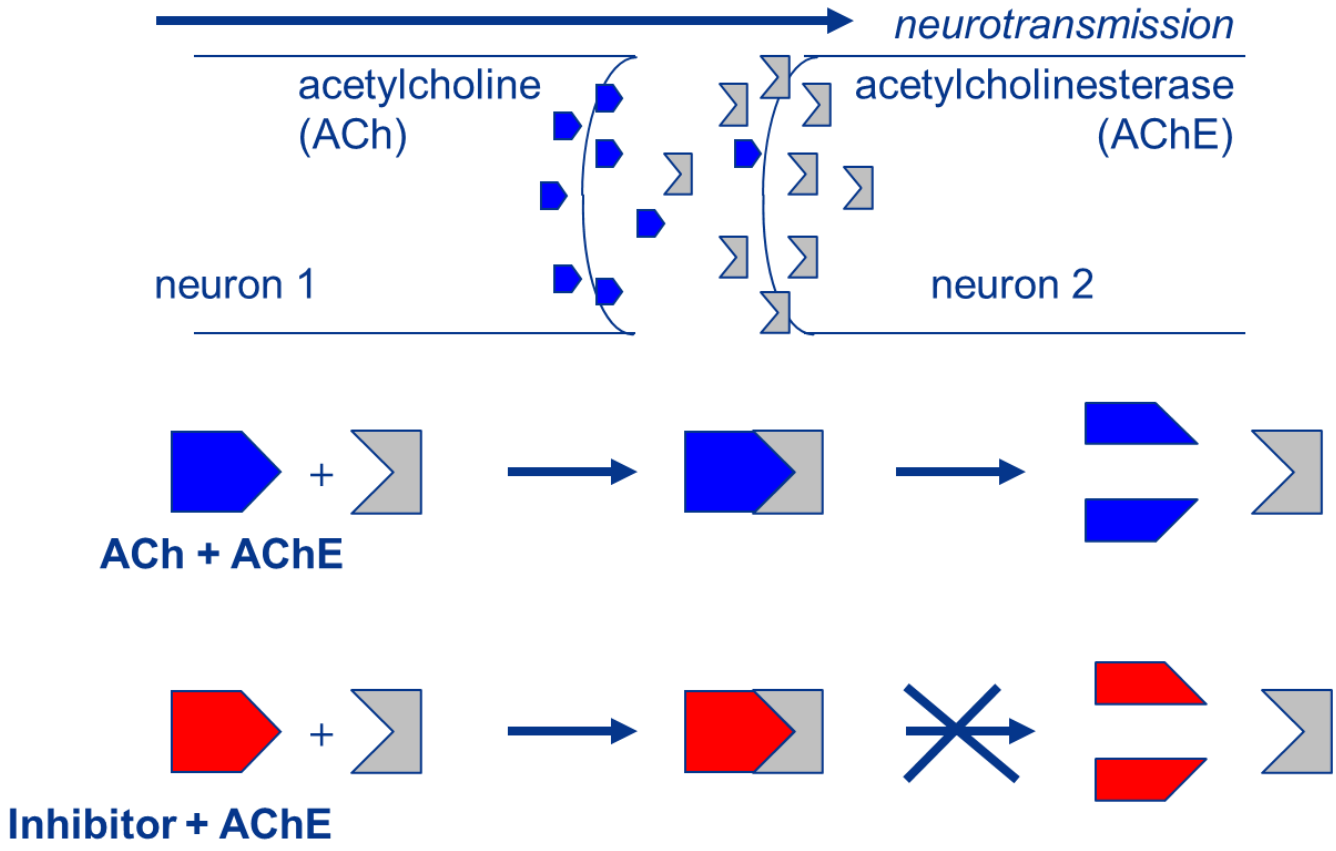


Figure 2: ACh (blue) is released from the presynaptic neuron into the synapse where it merges to and activates the AChR present on membrane of the postsynaptic cell (not shown). Meanwhile, AChE (grey) present in the synaptic cleft hydrolyses the ACh neurotransmitter to avoid overstimulation of the postsynaptic membrane. Organophosphate insecticides (red) bind to the AChE and prevent its reaction with ACh, causing accumulation of ACh.

Irreversible vs reversible enzyme inhibition

Organophosphate insecticides bind covalently to the AChE enzyme thereby causing irreversible enzyme inhibition. Irreversible enzyme inhibition progressively increases in time following first-order kinetics (link to section on Bioaccumulation and kinetic modelling). Recovery of enzyme activity can only be obtained by de novo synthesis of enzymes. In contrast to AChE inhibition, inhibition of the T4 transport function of TTR is reversible because the halogenated phenols bind to TTR in a non-covalent way. Similarly, non-covalent binding of a toxic compound to an enzyme causes reversible inhibition of the enzyme activity.

In addition to covalent and non-covalent enzyme binding, irreversible enzyme inhibition may occur when toxic compounds cause an error during enzyme synthesis. For instance, ions of essential metals, which are present as cofactors in the active site of many enzymes, may be replaced by ions of other metals during enzyme synthesis, yielding inactive enzymes. A classic example of such decreased enzyme activity is the inhibition of δ-aminolevulinic acid dehydratase (δ-ALAD) by lead. In this case, lead replaces zinc in the active site of the enzyme, thereby inhibiting a catalytic step in the synthesis of a precursor of heme, a cofactor of the protein hemoglobulin (link to section on Toxicity mechanisms of metals).

With respect to reversible enzyme inhibition, three types of inhibition can be distinguished, i.e. competitive, non-competitive, and uncompetitive inhibition (Figure 3).

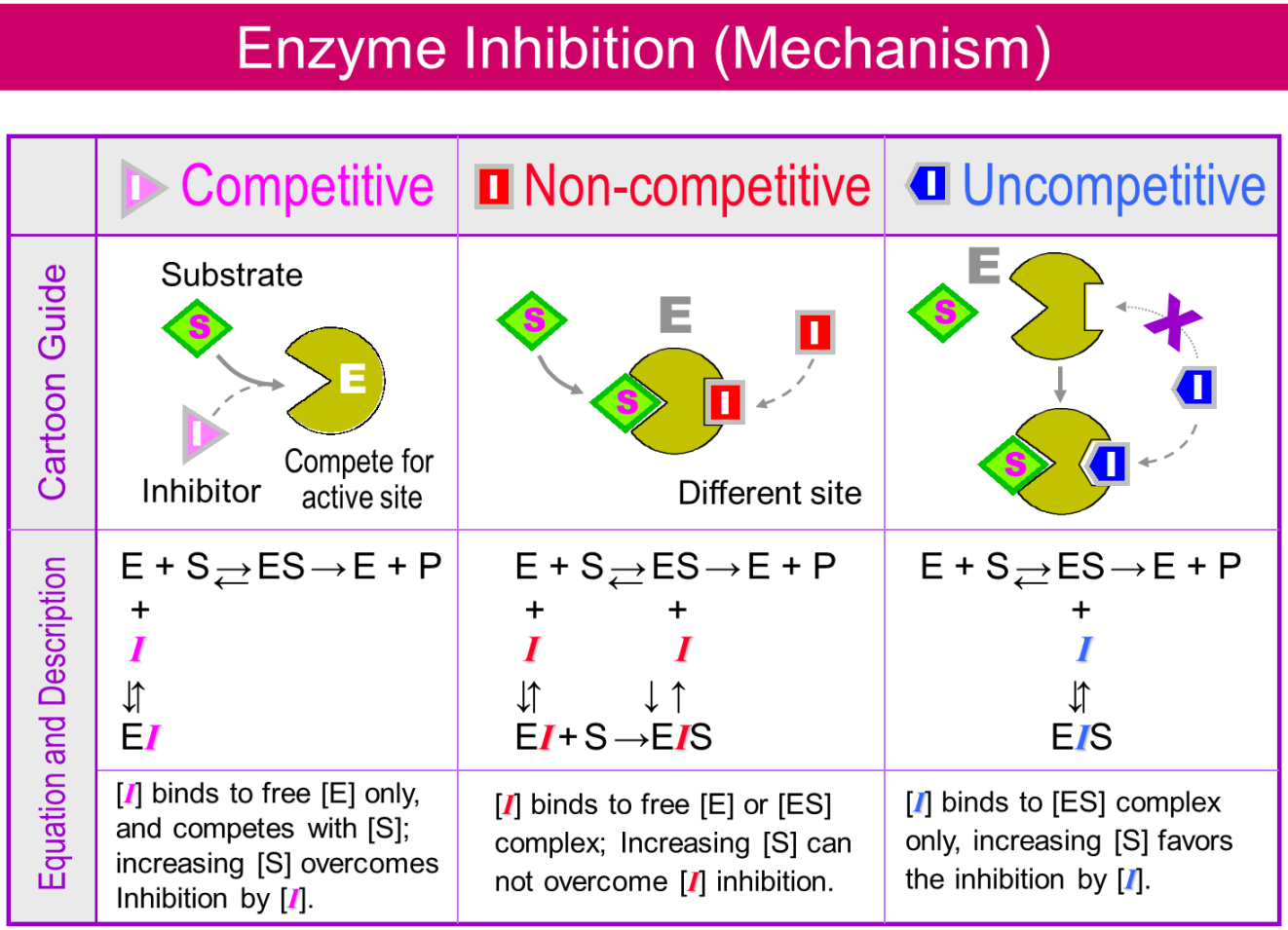


Figure 3: Three types of reversible enzyme inhibition, i.e. competitive (left), non-competitive (middle), and uncompetitive (right) binding. See text for further explanation. Source: juang.bst.ntu.edu.tw/files/Enz04%20inhibition.PPT

Competitive inhibition refers to a situation where the chemical competes ("fights") with the substrate for binding to the active site of the enzyme. Competitive inhibition is very specific, because it requires that the inhibitor resembles the substrate and fits in the same binding pocket of the active site. The TTR-binding example described above is a typical example of competitive inhibition between thyroid hormone and halogenated phenols for occupation of the TTR-binding site. A more classic example of competitive inhibition is the inhibition of beta-lactamase by penicillin. Beta-lactamase is an enzyme responsible for the hydrolysis of beta-

lactam, which is the final step in bacterial cell wall synthesis. By defective cell wall synthesis, penicillin is an antibiotic causing bacterial death.

Non-competitive inhibition refers to a situation where the chemical binds to an allosteric site of the enzyme (i.e. not the active site), thereby causing a conformational change of the active site. As a consequence, the substrate cannot enter the active site, or the active site becomes inactive, or the product cannot be released from the active site. For instance, echinocandin antifungal drugs non-competitively inhibit the enzyme 1,3-beta glucan synthase, which is responsible for the synthesis of beta-glucan, a major constituent of the fungal cell wall. Lack of beta-glucan in fungal cell walls prevents fungal resistance against osmotic forces, leading to cell lysis.

Uncompetitive inhibition refers to a situation where the chemical can only bind to the enzyme if the substrate is simultaneously bound. Substrate binding leads to a conformational change of the enzyme, which leads to the formation of an allosteric binding site for the inhibitor. Uncompetitive inhibition is more common in two-substrate enzyme reactions than in one-substrate enzyme reactions. An example of uncompetitive inhibition is the inhibition by lithium of the enzyme inositol mono phosphatase (IMPase), which is involved in recycling of the second messenger inositol-3-phospate (I3P) (link to section on Receptor interaction). IMPase is involved in the final step of dephosphorylating inositol monophosphate into inositol. Since lithium is the primary treatment for bipolar disorder, this observation has led to the inositol depletion hypothesis that inhibition of inositol phosphate metabolism offers a plausible explanation for the therapeutic effects of lithium.

4.2.1. Question 1

Explain how binding of organophosphate insecticides to acetylcholinesterase enzymes may cause neurotoxicity.

4.2.1. Question 2

Explain how organohalogenated phenols may cause decreased blood levels of thyroid hormone T4.

4.2.1. Question 3

What is the difference between a competitive and a non-competitive enzyme inhibitor?

4.2.1. Question 4

Is it possible to outcompete a competitive enzyme inhibitor by increasing the substrate concentration?

4.2.1. Question 5

Is it possible to outcompete a non-competitive enzyme inhibitor by increasing the substrate concentration?

4.2.2. Receptor interaction

Author: Timo Hamers

Reviewers: Frank van Belleghem and Ludek Blaha

Learning objectives

You should be able to

- explain the possible effects of compound interference with ion channels.
- explain the possible effects of compound interference with G-protein coupled receptors (GPCRs).
- explain the possible effects of compound interference with enzyme linked receptors.
- explain the possible effects of compound interference with nuclear receptors.
- understand what signalling pathways are and how they can be affected to toxic compounds

Key words: Ion channels, G-protein coupled receptors, enzyme linked receptors, nuclear receptors

Introduction

Receptor proteins specifically bind and respond to endogenous signalling ligands such as hormones, prostaglandins, growth factors, or neurotransmitters, by causing a typical cellular response. Receptor proteins can be located in the cell membrane, in the cytosol,

and in the nucleus of a cell. Agonistic receptor ligands activate the receptor protein whereas antagonistic ligands inactivate the receptor and prevent (endogenous) agonists from activating the receptor (Figure 1). Based on the role of the receptor protein, binding by ligands may interfere with:

1. ion channels
2. G-protein coupled receptors
3. enzyme-linked receptors
4. nuclear receptors.

Xenobiotic ligands can interfere with these cellular responses by acting as agonistic or antagonistic ligands.

.

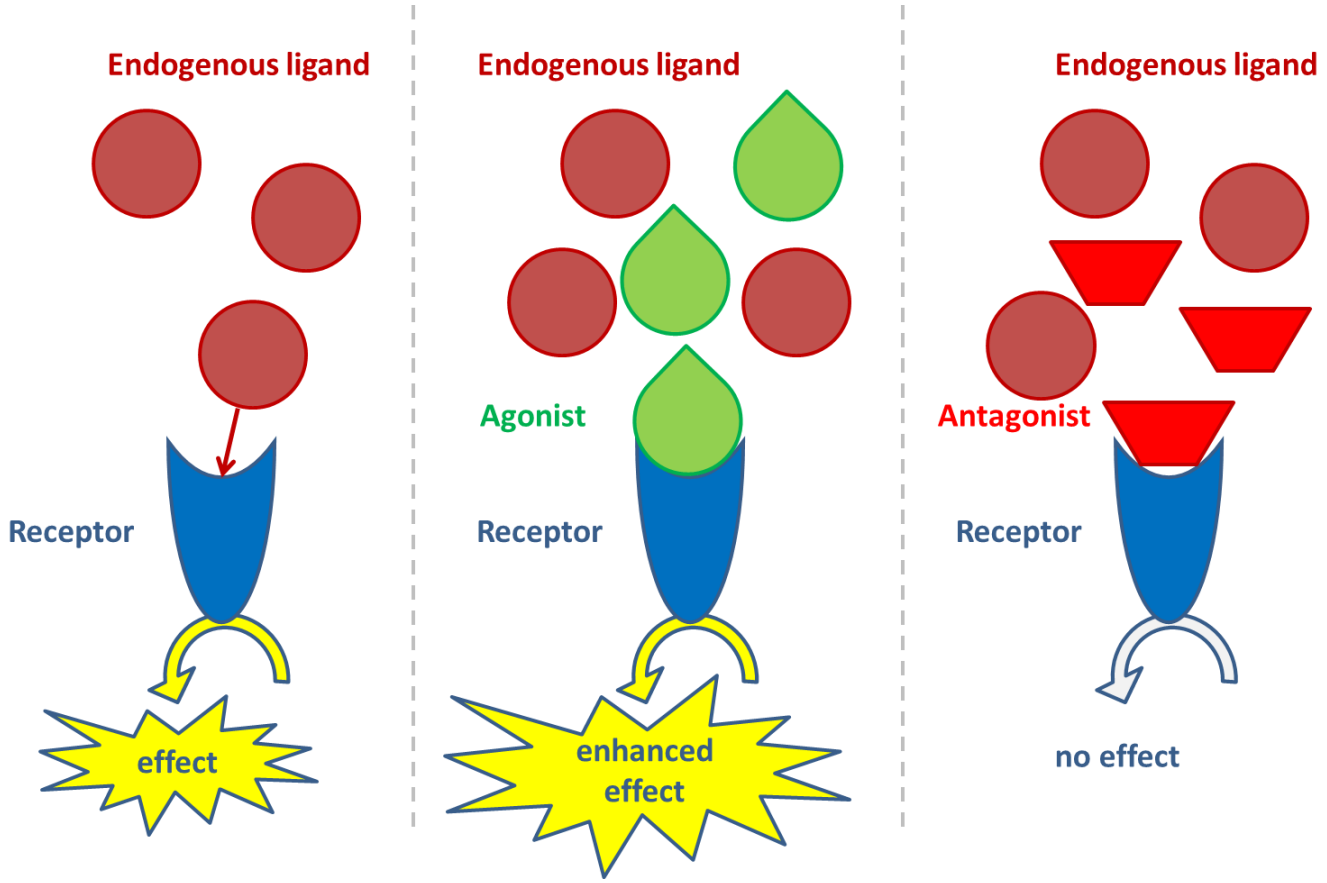


Figure 1: Activation by the endogenous ligand of a receptor leads to an effect. An agonistic compound may also activate the receptor and leads in cooperation with the endogenous ligand to an enhanced effect. An antagonistic compound also has binding affinity for the receptor, but cannot activate it. Instead, it prevents the endogenous ligand from binding, and activating the receptor, thereby preventing the effect.

1. Ion channels

Ion channels are transmembrane protein complexes that transport ions across a phospholipid bilayer membrane. Ion channels are especially important in neurotransmission, when stimulating neurotransmitters (e.g. acetylcholine or ACh) bind to the (so-called ionotropic) receptor part of the ion channel and open the ion channel for a very short (i.e. millisecond) period of time. As a result, ions can cross the membrane causing a change in transmembrane potential (see Figure). On the other hand, receptor-binding by inhibiting neurotransmitters (e.g. gamma-aminobutyric acid or GABA) prevents the opening of ion channels.

Compounds interfering with sodium channels, for instance, are neurotoxic compounds (see section on Neurotoxicity). They can either block the ion channels or keep them in a prolonged or permanently open state. Many compounds known to interfere with ion channels are natural toxins. For instance, tetrodotoxin (TTX), which is produced by marine bacteria and highly accumulated in

puffer fish, and saxitoxin, which is produced by dinoflagellates and is accumulated in shellfish are capable of blocking voltage-gated sodium channels in nerve cells. In contrast, ciguatoxin, which is another persistent toxin produced by dinoflagellates that accumulates in predatory fish positioned high in the food chain, causes prolongation of the opening of voltage-gated sodium channels. Some pesticides like DDT and pyrethroid insecticides also prevent closure of voltage-gated sodium channels in nerve cells. As a consequence, full repolarization of the membrane potential is not achieved. As a consequence, the nerve cells do not reach the resting potential and any new stimulus that would be too low to reach the threshold for depolarization under normal conditions, will now cause a new action potential. In other words, the nerve cells become hyperexcitable and undergo a series of action potentials (repetitive firing) causing tremors and hyperthermia.

2. G-protein coupled receptors (GPCRs)

GPCRs are transmembrane receptors that transfer an extracellular signal into an activated G-protein that is connected to the receptor on the intracellular side of the membrane. G-proteins are heterotrimer proteins consisting of three subunits alpha, beta, and gamma, of which the alpha subunit - in inactivated form - contains a guanosine diphosphate (GDP) molecule. Upon binding by endogenous ligands such as hormones, prostaglandins, or neurotransmitters (i.e. the signal or "first messenger") to the (so-called metabotropic) receptor, a conformational change in the GPCR complex leads to an exchange of the GDP for a guanosine triphosphate (GTP) molecule in the alpha monomer part of the G-protein, causing release of the activated alpha subunit from the beta/gamma dimer part. The activated alpha monomer can interact with several target enzymes causing an increase in "second messengers" starting signal transduction pathways (see point 3 Enzyme-linked receptors). The remaining beta-gamma complex may also move along the inner membrane surface and affect the activity of other proteins (Figure 2).

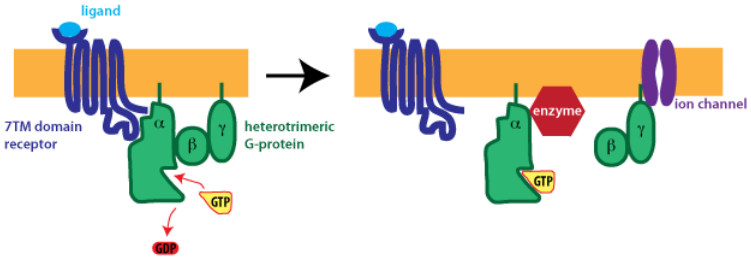


Figure 2: Mechanism of GPCR-activation: ligand binding causes a conformational change leading to the release of an activated alpha monomer, which interacts with a target enzyme (causing an increase of second messengers), and a beta-gamma dimer, which may directly affect activity of other proteins (e.g. an ion channel). Source: courses.washington.edu/conj/bess/gpcr/gpcr.htm

Two major enzymes that are activated by the alpha monomer are adenylyl cyclase causing an increase in second messenger cyclic AMP (cAMP) and phospholipase C causing an increase in second messenger diacylglycerol (DAG). In turn, cAMP and DAG activate protein kinases, which can phosphorylate many other enzymes. Activated phospholipase C also causes an increase in levels of the second messenger inositol-3-phosphate (I3P), which opens ion channels in the endoplasmic reticulum causing a release of calcium from the endoplasmic store, which also acts as a second messenger. On the other hand, the increase in cytosolic calcium levels is simultaneously tempered by the beta/gamma dimer, which can inhibit voltage-gated calcium channels in the cell membrane. Ultimately, the GPCR signal is extinguished by slow dephosphorylation of GTP into GDP by the activated alpha monomer, causing it to rearrange with the beta/gamma dimer into the original inactivated trimer G-protein (see also courses.washington.edu/conj/bess/gpcr/gpcr.htm).

The most well-known example of disruption of GPCR signalling is by cholera toxin (see text block Cholera toxin below).

Despite the recognized importance of GPRCs in medicine and pharmacology, little attention has so-far been paid in toxicology to interaction of xenobiotics with GPCRs. Although a limited number of studies have demonstrated that endocrine disrupting compounds including PAHs, dioxins, phthalates, bisphenol-A, and DDT can interact with GPCR signalling, the toxicological

implications of these interactions (especially with respect to disturbed energetic metabolism) remain subject for further research (see review by Le Ferrec and Øvrevik, 2018).

Cholera toxin

Cholera toxin is a so-called AB exotoxin by *Vibrio cholerae* bacteria, consisting of an "active" A-part and a "binding" B-part (see http://www.sumanasinc.com/webcontent/animations/content/diphtheria.html). Upon binding by the B-part to the intestinal epithelium membrane, the entire AB complex is internalized into the cell via endocytosis, and the active A-part is released. This A-part adds an ADP-ribose group to G-proteins making the GTP dephosphorylation of activated G-proteins impossible. As a consequence, activated G-proteins remain in a permanent active state, adenylyl cyclase is permanently activated and cAMP levels rise, which in turn cause an imbalance in ion housekeeping, i.e. an excessive secretion of chloride ions to the gut lumen and a decreased uptake of sodium ions from the gut lumen. Due to the increased osmotic pressure, water is released to the gut lumen causing dehydration and severe diarrhoea ("rice-water stool").

3. Enzyme-linked receptors

Enzyme-linked receptors are transmembrane receptors that transfer an extracellular signal into an intracellular enzymatic activity. Most enzyme-linked receptors belong to the family of receptor tyrosine kinase (RTK) proteins. Upon binding by endogenous ligands such as hormones, cytokines, or growth factors (i.e. the signal or primary messenger) to the extracellular domain of the receptors, the receptor monomers dimerize and develop kinase activity, i.e. become capable of coupling of a phosphate group donated by a high-energy donor molecule to an acceptor protein. The first substrate for this phosphorylation activity is the dimerized receptor itself, which accepts a phosphate group donated by ATP on its intracellular tyrosine residues. This autophosphorylation is the first step of a signalling pathway consisting of a cascade of subsequent phosphorylation steps of other kinase proteins (i.e. signal transduction), ultimately leading to transcriptional activation of genes followed by a cellular response (Figure 3).

Figure in preparation

Figure 3: *Upon ligand binding, tyrosine kinase receptor (TKR) proteins become autophosphorylated and may phosphorylate (i.e. activate other proteins), including other kinases.*

Xenobiotic compounds can interfere with these signalling pathways in many different ways. Compounds may avoid binding of the endogenous ligand, by blocking the receptor or by chelating the endogenous ligands. Most RTK inhibitors inhibit the kinase activity directly by acting as a competitive inhibitor for ATP binding to the tyrosine residues. Many RTK inhibitors are used in cancer treatment, because RTK overactivity is typical for many types of cancer. This overactivity may for instance be caused by increased levels of receptor-activating growth factors, or to spontaneous dimerization when the receptor is overexpressed or mutated).

4. Nuclear receptors

Nuclear receptors are proteins that are activated by endogenous compounds (often hormones) leading ultimately to expression of genes specifically regulated by these receptors. Apart from ligand binding, activation of most nuclear receptors requires dimerization with a coactivating transcription factor. While some nuclear receptors are located in the nucleus in inactive form (e.g. the thyroid hormone receptor), most nuclear receptors are located in the cytosol, where they are bound to co-repressor proteins (often heat-shock proteins) keeping them in an inactive state. Upon ligand binding to the ligand binding domain (LBD) of the receptor, the co-repressor proteins are released and the receptor either forms a homodimer with a similar activated nuclear receptor or forms a heterodimer with a different nuclear receptor, which is often the retinoid-X receptor (RXR) for nuclear hormone receptors. Before or after dimerization, activated nuclear receptors are translocated to the nucleus. In the nucleus, they bind through their DNA-binding domain (DBD, or "zinc finger") to a responsive element in the DNA located in the promotor region of receptor-responsive genes. Consequently, these genes are transcribed to mRNA in the nucleus, which is further translated into proteins in the cell cytoplasm, see Figure 4).

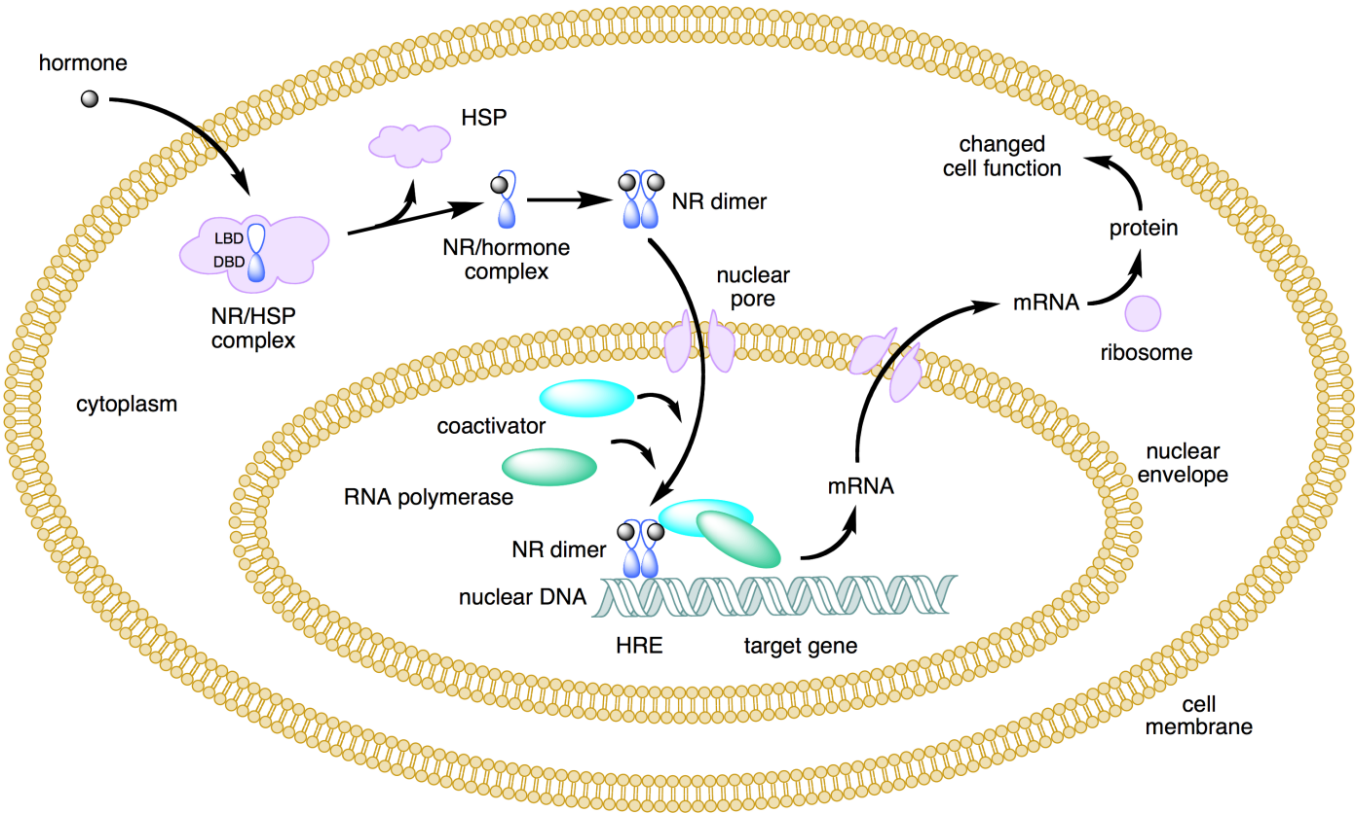


Figure 4: Activation of a cytosolic nuclear receptor (NR). Upon ligand binding (e.g. a hormone), the heat shock proteins (HSP) dissociate from the ligand-receptor complex, which forms a heterodimer before entering the nucleus. After recruiting other coactivating transcription factors, the activated dimer binds to the hormone response element (HRE). RNA polymerase binds to this complex and starts transcription of mRNA, which is excreted from the nucleus into the cytosol and transcribed in corresponding proteins. . Source: https://upload.wikimedia.org/Wikipedia/commons/3/3f/Nuclear_receptor_action.png

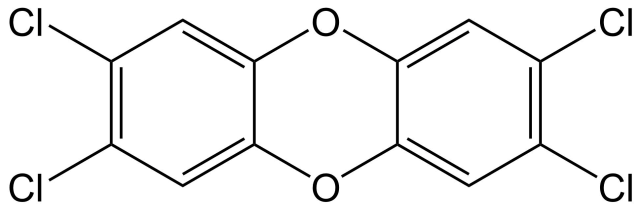
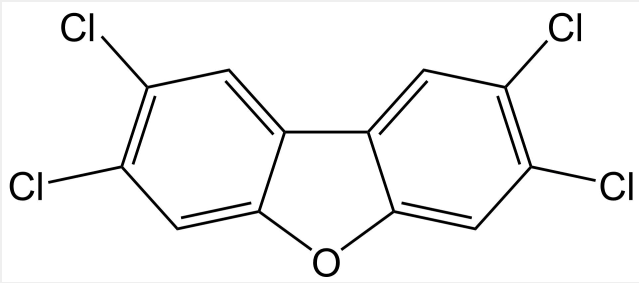
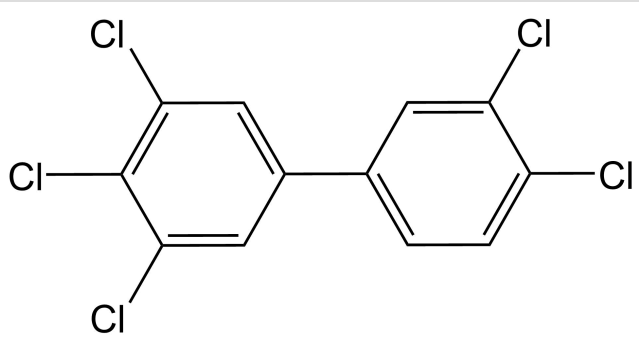
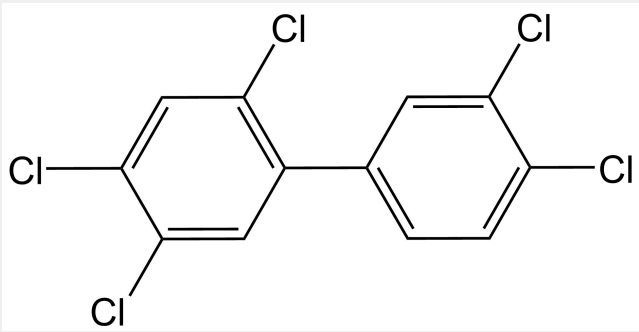
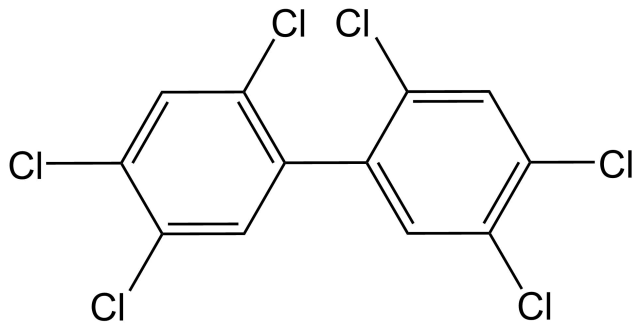
Xenobiotic compounds may act as agonist or antagonists of nuclear receptor activation. Chemicals that act as a nuclear receptor agonist mimic the action of the endogenous activator(s), whereas chemicals that act as a nuclear receptor antagonist basically block the LBD of the receptor, preventing the binding of the endogenous activator(s). Over the past decades, interaction of xenobiotics with nuclear receptors involved in signalling of both steroid and non-steroid hormones has gained a lot of attention of researchers

investigating endocrine disruption (link to section on Endocrine Disruption). Nuclear receptor activation is also the key mechanism in dioxin-like toxicity (see text block dioxin-like toxicity below).

Dioxin-like toxicity

The term dioxins refers to polyhalogenated dibenzo-[p]-dioxin (PHDD) compounds, which are planar molecules consisting of two halogenated aromatic rings, which are connected by two ether bridges. The most potent and well-studied dioxin is 2,3,7,8-tetrachloro-[p]-dibenzodioxin (2,3,7,8-TCDD), which is often too simply referred to as TCDD or even just "dioxin". Other compounds with similar properties (dioxin-like compounds) include polyhalogenated dibenzo-[p]-furan (PHDF) compounds (often too simply referred to as "furans"), which are planar molecules consisting of two halogenated aromatic rings connected by one ether bridge and one carbon-carbon bond. A third major class of dioxin-like compounds belong to the polyhalogenated biphenyls (PHB), which consist of two halogenated aromatic rings connected only by a carbon-carbon bond. The most well-known compounds belonging to this latter category are the polychlorinated biphenyls (PCBs). Of all PHDD, PHDF or PHB compounds, only the persistent and planar compounds are considered dioxin-like compounds. For the PHBs, this implies that they should contain zero or at maximum one halogen-substitution in

any of the four ortho-positions (see examples below). Non-ortho-substituted PHBs can easily obtain a planar confirmation with the two aromatic rings in one planar field, whereas mono-ortho-substituted PHBs can obtain such confirmation at higher energetic costs..

	2,3,7,8-tetrachlorodibenzo-[p]-dioxin (2,3,7,8-TCDD) is the most potent and well-studied dioxin-like compound, usually too simply referred to as "dioxin".
	2,3,7,8-tetrachlorodibenzo-[p]-furan (2,3,7,8-TCDF) a dioxin-like compound equally potent to 2,3,7,8-TCDD. It is usually too simply referred to as "furan".
	3,3',4,4',5-pentachlorinated biphenyl (PCB-126) is the most potent dioxin-like PCB compound, with no chlorine substitution in any of the four ortho positions next to the carbon-carbon bridge
	2,3',4,4',5-pentachlorinated biphenyl (PCB-118) is a weak dioxin-like PCB compound, with one chlorine substitution in the four ortho positions next to the carbon-carbon bridge
	2,2',4,4',5,5'-hexachlorinated biphenyl (PCB-153) is a non-dioxin-like (NDL) PCB compound, with two chlorine substitution in the four ortho positions next to the carbon-carbon bridge

The planar composition is required for the dioxin-like compounds to fit as a key in the lock of the arylhydrocarbon (AhR) receptor (also known as the "dioxin-receptor or DR), present in the cytosol. The activated AhR then dissociates from its repressor proteins, is

translocated to the nucleus, and forms a heterodimer with the AhR nuclear translocator (ARNT). The AhR-ARNT complex binds to dioxin-response elements (DRE) in the promotor regions of dioxin-responsive genes in the DNA, ultimately leading to transcription and translation of these genes (see Figure 1 in Denison & Nagy, 2003). Famous examples of such genes belong to the CYP1, UGT, and GST families, which are Phase I and Phase II metabolic enzymes whose activation by the AhR-ARNT complex is a natural response triggered by the need to remove xenobiotics (link to section on Xenobiotic metabolism and defence). Other genes with a DRE in their promotor region include genes involved in protein phosphorylation, such as the proto-oncogen c-raf and the cyclin dependent kinase inhibitor p27.

This classical mechanism of ligand:AhR:ARNT:DRE complex-dependent induction of gene expression, however, cannot explain all the different types of toxicity observed for dioxins, including immunotoxicity, reproductive toxicity and developmental toxicity. Still, these effects are known to be mediated through the AhR as well, as they were not observed in AhR knockout mice. This can partly be explained by the fact that not all genes that are under transcriptional control of a DRE are known yet. Moreover, AhR dependent mechanisms other than this classical mechanism have been described. For instance, AhR activation may have anti-estrogenic effects because activated AhR (1) binds to the estrogen receptor (ER) and targets it for degradation, (2) binds (with ARNT) to inhibitory DREs in the promotor of ER-dependent genes, and (3) competes with the ER-dimer for common coactivators. Although dioxin-like compounds absolutely require the AhR to exert their major toxicological effects, several AhR independent effects have been described as well, such as AhR-independent alterations in gene expression and changes in Ca^{2+} influx related to changes in protein kinase activity.

Apart from the persistent halogenated dioxinlike compounds described above, other compounds may also activate the AhR, including natural AhR agonists (nAhRAs) found in food (e.g. indolo[3,2-b]carbazole (ICZ) in cruciferous vegetables, bergamottin in grapefruits, tangeretin in citrus fruits), and other planar aromatic compounds, including polycyclic aromatic hydrocarbons (PAHs) produced by incomplete combustion of organic fuels. Upon activation of the AhR, these non-persistent compounds are metabolized by the induced CYP1A biotransformation enzymes. In addition, an endogenous AhR ligand called 6-formylindolo[3,2-b]carbazole (FICZ) has been identified. FICZ is a mediator in many physiological processes, including immune responses, cell growth and differentiation. Endogenous FICZ levels are regulated by a negative feedback FICZ/AhR/CYP1A loop, i.e. FICZ activates AhR and is metabolized by the subsequently induced CYP1A. Dysregulation of this negative feedback loop by other AhR agonists may disrupt FICZ functioning, and could possibly explain some of the effects observed for dioxinlike compounds.

Further reading:

Denison, M.S., Soshilov, A.A., He, G., De Groot, D.E., Zhao, B. (2011). Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the Aryl hydrocarbon (Dioxin) Receptor. Toxicological Sciences 124, 1-22.

Further reading:

Boelsterli, U.A. (2009). *Mechanistic Toxicology* (2nd edition). Informa Healthcare, New York, London.

Le Ferrec, E., Øvrevik J. (2018). G-protein coupled receptors (GPCR) and environmental exposure. Consequences for cell metabolism using the b-adrenoceptors as example. Current Opinion in Toxicology 8, 14-19.

courses.washington.edu/conj/bess/gpcr/gpcr.htm

4.2.2. Question 1

Why are compounds interfering with ion channels mainly neurotoxic compounds?

4.2.2. Question 2

GPRC signalling is not only disrupted through interaction with the receptor. What alternative mechanisms can play a role?

4.2.2. Question 3

What is the main effect of activating enzyme-linked receptors?

4.2.2. Question 4

What happens if a compound binds to a nuclear receptor?

4.2.3. Oxidative stress - I.

Reactive oxygen species and antioxidants

Author: Frank van Belleghem

Reviewers: Raymond Niesink, Kees van Gestel, Éva Hideg

Learning objectives:

You should be able to

- explain what oxidative stress is, under which circumstances it arises and why it is important in toxicology.
- describe what reactive oxygen species are and how they are produced.
- describe how levels of reactive oxygen species are kept under control.
- make a distinction between enzymatic and non-enzymatic antioxidants.

Keywords: Reactive oxygen species, Fenton reaction, Enzymatic antioxidants, Non-enzymatic antioxidants, Lipid peroxidation.

Reactive oxygen species

Molecular oxygen (O_2) is a byproduct of photosynthesis and essential to all heterotrophic cells because it functions as the terminal electron acceptor during the oxidation of organic substances in aerobic respiration. This process results in the reduction of O_2 to water, leading to the formation of chemical energy and reducing power. The reason why O_2 can be reduced with relative ease in biological systems can be found in the physicochemical properties of the oxygen molecule (in the triplet ground state, *i.e.* as it occurs in the atmosphere). Because of its electron configuration, O_2 is actually a biradical that can act as an electron acceptor. The outer molecular orbitals of O_2 each contain one electron, the spins of these electrons are parallel (Figure 1). As a result, oxygen (in the ground state) is not very reactive because, according to the Pauli exclusion principle, only one electron at a time can react with other electrons in a covalent bond. As a consequence, oxygen can only undergo univalent reductions, and the complete reduction of oxygen to water requires the sequential addition of four electrons leading to the formation of one-, two-, three-electron oxygen intermediates (Figure 1). These oxygen intermediates are, in sequence, the ***superoxide anion radical*** **($O_2^{\bullet -}$)**, ***hydrogen peroxide*** ***(H_2O_2)*** and the ***hydroxyl radical*** ***($^{\bullet}OH$)***.

Another reactive oxygen species of importance is ***singlet oxygen*** ***($^1O_2$ or $^1\Delta_g$)***. Singlet oxygen is formed by converting ground-state molecular oxygen into an excited energy state, which is much more reactive than the normal ground-state molecular oxygen. Singlet oxygen is typically generated by a process called photosensitization, for example in the lens of the eye. Photosensitization occurs when light (UV) absorption by an endogenous or xenobiotic substance lifts the compound to a higher energy state (a high-energy triplet intermediate) which can transfer its energy to oxygen, forming highly reactive singlet oxygen. Apart from oxygen-dependent photodynamic reactions, singlet oxygen is also produced by neutrophils and this has been suggested to be important for bacterial killing through the formation of ozone (O_3) (Onyango, 2016).

Because these oxygen intermediates are potentially deleterious products that can damage cellular components, they are referred to as ***reactive oxygen species (ROS)***. ROS are also often termed 'free radicals' but this is incorrect because not all ROS are radicals (e.g. H_2O_2, $^1O_2$ and O_3). Moreover, as all radicals are (currently) considered as unattached, the prefix 'free' is actually unnecessary (Koppenol & Traynham, 1996).

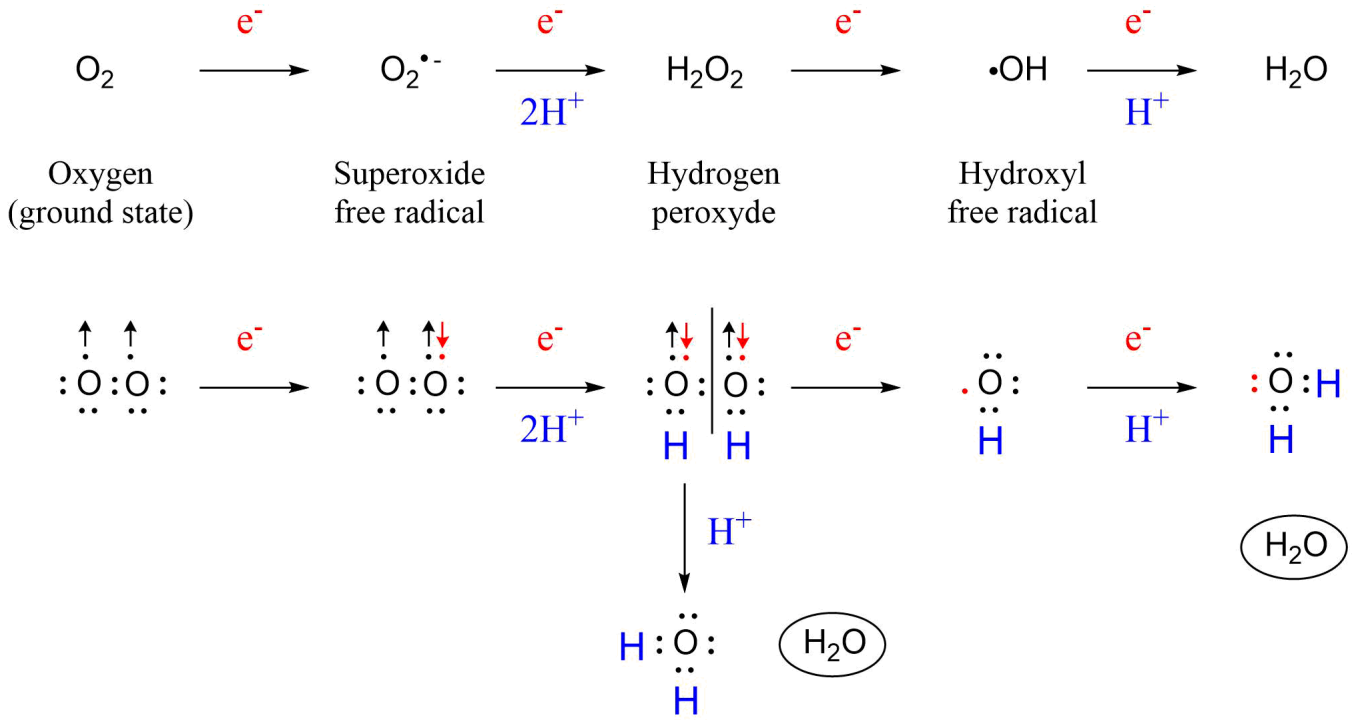


Figure adapted from Edreva (2005) by Steven Droge.

ROS are byproducts of aerobic metabolism in the different organelles of cells, for instance respiration or photosynthesis, or as part of defenses against pathogens. Endogenous sources of reactive oxygen species include oxidative phosphorylation, P450 metabolism, peroxisomes and inflammatory cell activation. For example, superoxide anion radicals are endogenously formed from the reduction of oxygen by the semiquinone of ubiquinone (coenzyme Q), a coenzyme widely distributed in plants, animals, and microorganisms. Ubiquinones function in conjunction with enzymes in cellular respiration (*i.e.*, oxidation-reduction processes). The superoxide anion radical is formed when one electron is taken up by one of the antibonding π*-orbitals (formed by two 2p atomic orbitals) of molecular oxygen.

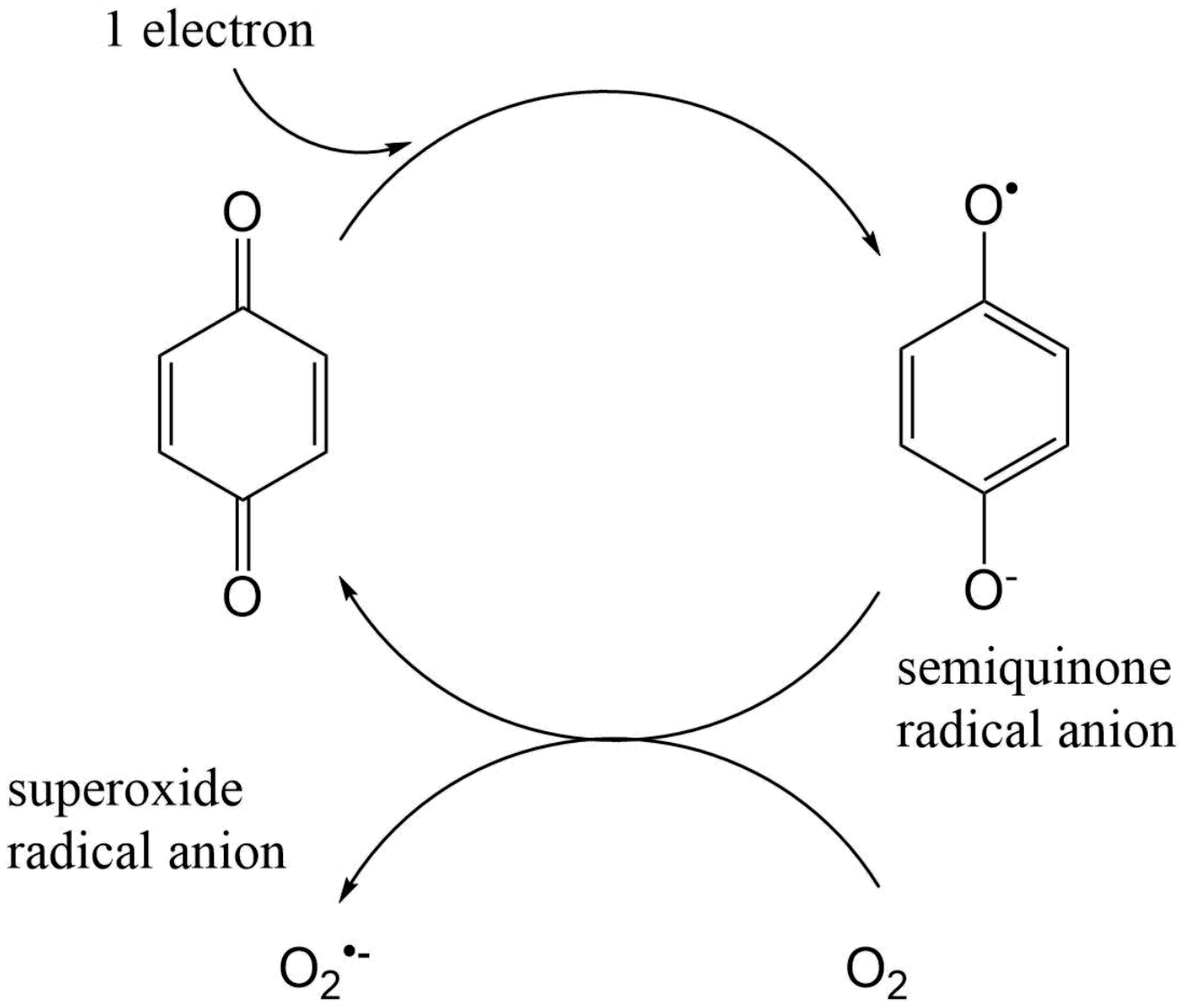


Figure adapted from Bolton & Dunlap (2016) by Steven Droge.

A second example of an endogenous source of superoxide anion radicals is the auto-oxidation of reduced heme proteins. It is known, for example, that oxyferrocytochrome P-450 substrate complexes may undergo auto-oxidation and subsequently split into (ferri) cytochrome P-450, a superoxide anion radical and the substrate (S). This process is known as the uncoupling of the cytochrome P-450 (CYP) cycle and also referred to as the oxidase activity of cytochrome P-450. However, it should be mentioned that this is not the normal functioning of CYP. Only when the transfer of an oxygen atom to a substrate is not tightly coupled to NADPH utilization, so that electrons derived from NADPH are transferred to oxygen to produce $O_2^{\bullet -}$ (and also H_2O_2).

$$\text{P-450-Fe}^{2+}\text{-- O}_2 \;(\text{S}) \rightleftharpoons \text{P-450-Fe}^{3+}\text{-- O}_2^{\bullet -}\;(\text{S}) \longrightarrow \text{P-450-Fe}^{3+} + \text{O}_2^{\bullet -} + \text{S}$$

Table 1 shows the key oxygen species and their biological half-life, their migration distance, the endogenous source and their reaction with biological compounds.

Table 1. The key oxygen species and their characteristics (table adapted from Das & Roychoudhury, 2014)

ROS species	Half-life ($T_{1/2}$)	Migration distance	Endogenous source	Mode of action
Superoxide anion radical ($O_2^{\bullet -}$)	1-4 µs	30 nm	Mitochondria, cytochrome P450, macrophage/ inflammatory cells, membranes, chloroplasts	Reacts with compounds with double bonds
Hydroxyl radical ($^{\bullet}OH$)	1 µs	1 nm	Mitochondria, membranes, chloroplasts	Reacts vigorously with all biomolecules.
Hydrogen peroxide (H_2O_2)	1 ms	1 µm	Mitochondria, membranes, peroxisomes, chloroplasts	Oxidizes proteins by reacting with the Cys residue.
Singlet Oxygen	1-4 µs	30 nm	Mitochondria, membranes, chloroplasts	Oxidizes proteins, polyunsaturated fatty acids and DNA

Because of their reactivity, at elevated levels ROS can indiscriminately damage cellular components such as lipids, proteins and nucleic acids. In particular the superoxide anion radical and hydroxyl radicals that possess an unpaired electron are very reactive. In fact, hydroxyl has the highest 1-electron reduction potential, making it the single most reactive radical known. Hydroxyl radicals (Figure 1) can arise from hydrogen peroxide in the presence of redox-active transition metal, notably $Fe^{2+/3+}$ or $Cu^{+/2+}$, via the ***Fenton reaction***. In case of iron, for this reaction to take place, the oxidized form (Fe^{3+}) has to be reduced to Fe^{2+}. This means that Fe^{2+} is only released in an acidic environment (local hypoxia) or in the presence of superoxide anion radicals. The reduction of Fe^{3+}, followed by the interaction with hydrogen peroxide, leading to the generation of hydroxyl radical, is called the iron catalyzed ***Haber-Weiss reaction***.

$$O_2^{\bullet -} + Fe^{3+} \longrightarrow O_2 + Fe^{2+}$$

Fenton $\quad H_2O_2 + Fe^{2+} \longrightarrow OH^- + {}^{\bullet}OH + Fe^{3+}$

Haber-Weiss $\quad O_2^{\bullet -} + H_2O_2 \longrightarrow OH^- + O_2 + {}^{\bullet}OH$



Keeping reactive oxygen species under control

In order to keep the ROS concentrations at low physiologic levels, aerobic organisms have evolved complex antioxidant defense systems that include antioxidant components that are enzymatic and non-enzymatic. These are cellular mechanisms that are evolved to inhibit oxidation by quenching ROS. Three classes of ***enzymes*** are known to provide protection against reactive oxygen species: the ***superoxide dismutases*** that catalyze the dismutation of the superoxide anion radical, and the ***catalases*** and ***peroxidases*** that react specifically with hydrogen peroxide. These antioxidant enzymes can be seen as a ***first-line defense*** as they prevent the conversion of the less reactive oxygen species, superoxide anion radical and hydrogen peroxide, to more reactive species such as the hydroxyl radical. **The *second line of defense*** largely consists of ***non-enzymatic substances*** that eliminate radicals such as ***glutathione*** and ***vitamins E*** and ***C***. An overview of the cellular defense system is provided in Figure 3.

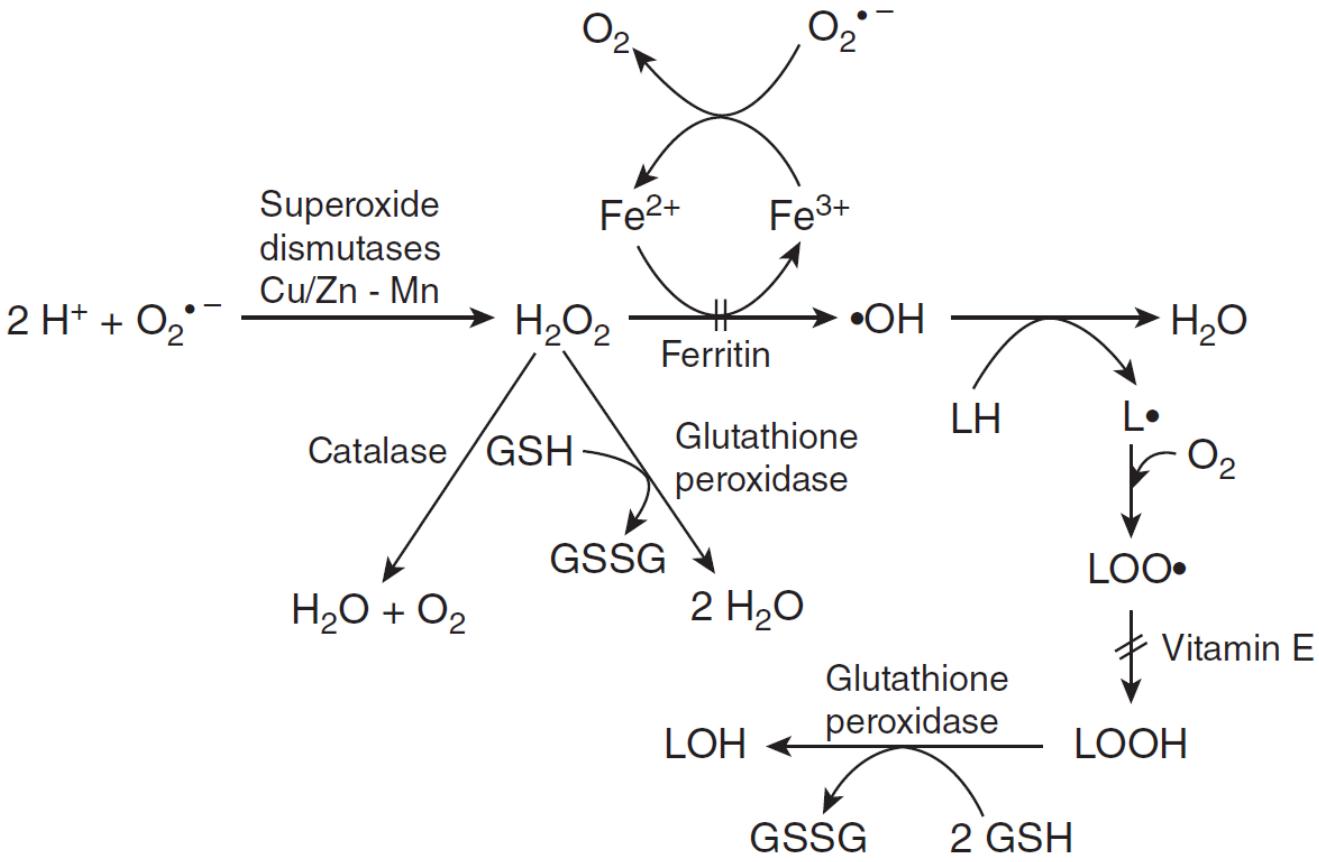


Figure adapted from Smart & Hodgson (2018) by Steven Droge.

Enzymatic antioxidants

Superoxide dismutases **(SODs)** are metal-containing proteins (metalloenzymes) that catalyze the dismutation of the superoxide anion radical to molecular oxygen in the ground state and hydrogen peroxide, as illustrated by following reactions:

a $O_2^{\bullet-} + SOD\text{-}Cu^{2+} \longrightarrow O_2 + SOD\text{-}Cu^{+}$

b $O_2^{\bullet-} + SOD\text{-}Cu^{2+} + 2\,H^{+} \longrightarrow H_2O_2 + SOD\text{-}Cu^{+}$

Dismutation of superoxide anion radicals acts in the first part of the reaction with the superoxide anion radical as a reducing agent (a), and as an oxidant in the second part (b). Different types of SOD are located in different cellular locations, for instance Cu-Zn-SOD are mainly located in the cytosol of eukaryotes, Mn-SOD in mitochondria and prokaryotes, Fe-SOD in chloroplasts and prokaryotes and Ni-SOD in prokaryotes. Mn, Fe, Cu and Ni are the redox active metals in the enzymes, whereas Zn not being catalytic in the Cu-Zn-SOD.

H_2O_2 is further degraded by catalase and peroxidase. ***Catalase (CAT)*** contains four iron-containing heme groups that allow the enzyme to react with the hydrogen peroxide and is usually located in peroxisomes, which are organelles with a high rate of ROS production. Catalase converts hydrogen peroxide to water and oxygen. In fact, catalase cooperates with superoxide dismutase in the removal of the hydrogen peroxide resulting from the dismutation reaction. Catalase acts only on hydrogen peroxide, not on organic hydroperoxide.

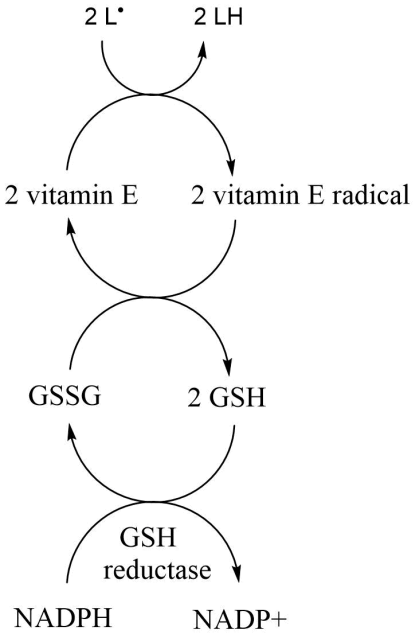
Peroxidases (Px) are hemoproteins that utilize H_2O_2 to oxidize a variety of endogenous and exogenous substrates. An important peroxidase enzyme family is the selenium-cysteine containing ***Glutathione peroxidase (GPx)***, present in the cytosol and

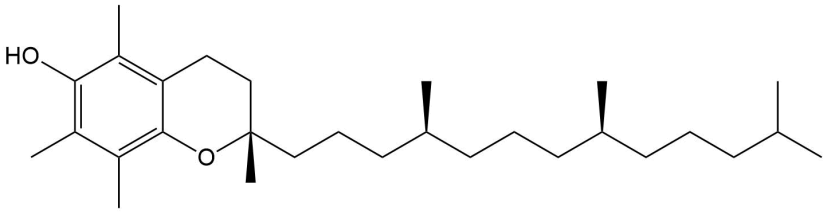
mitochondria. It catalyzes the conversion of hydrogen H_2O_2 to H_2O via the oxidation of reduced glutathione (GSH) into its disulfide form glutathione disulfide (GSSG). Glutathione peroxidase catalyzes not only the conversion of hydrogen peroxide, but also that of organic peroxides. It can transform various peroxides, e.g. the hydroperoxides of lipids. Glutathione peroxidase is found in both the cytosol and in the mitochondria. In the cytosol, the enzyme is present in special vesicles.
Another group of enzymes, not further described here, are the ***Peroxiredoxins (Prxs)***, present in the cytosol, mitochondria, and endoplasmic reticulum, use a pair of cysteine residues to reduce and thereby detoxify hydrogen peroxide and other peroxides. It has to be mentioned that no enzymes react with hydroxyl radical or singlet oxygen.

Non-enzymatic antioxidants

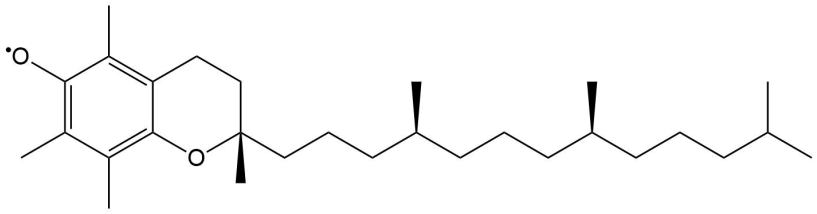
The *second line of defense* largely consists of ***non-enzymatic substances*** that eliminate radicals. The major antioxidant is ***glutathione (GSH)***, which acts as a nucleophilic scavenger of toxic compounds, trapping electrophilic metabolites by forming a thioether bond between the cysteine residue of GSH and the electrophile. The result generally is a less reactive and more water-soluble conjugate that can easily be excreted (see also phase II biotransformation reactions). GSH also is a co-substrate for the enzymatic (GS peroxidase-catalyzed) degradation of H_2O_2 and it keeps cells in a reduced state and is involved in the regeneration of oxidized proteins.

Other important radical scavengers of the cell are the vitamins E and C. ***Vitamin E* (α-tocopherol)** is lipophilic and is incorporated in cell membranes and subcellular organelles (endoplasmic reticulum , mitochondria, cell nuclei) and reacts with lipid peroxides. α-Tocopherol can be divided into two parts, a lipophilic phytyl tail (intercalating with fatty acid residues of phospholipids) and a more hydrophilic chroman head with a phenolic group (facing the cytoplasm). This phenolic group can reduce radicals (e.g. lipid peroxy radicals ($LOO^\bullet$, see Figure 2, for explanation of lipid peroxidation, see section on Oxidative stress II: induction by chemical exposure and possible effects) and is thereby oxidized in turn to the tocopheryl radical which is relatively unreactive because it is stabilized by resonance. The radical is regenerated by vitamin C or by reduced glutathione (Figure 4). Oxidized non-enzymatic antioxidants are regenerated by various enzymes such as glutathione.





α-Tocopherol (vitamin E)



vitamin E radical

Figure adapted from Niesink et al. (1996) by Steven Droge.

Vitamin C (ascorbic acid) is a water-soluble antioxidant and is present in the cytoplasm. Ascorbic acid is an electron donor which reacts quite rapidly with the superoxide anion radical and peroxyl radicals, but is generally ineffective in detoxifying hydroxyl radicals because of its extreme reactivity it does not reach the antioxidant (See Klaassen, 2013). Moreover, it regenerates α-Tocopherol in combination with reduced GSH or compounds capable of donating reducing equivalents (Nimse and Pal, 2015): Figure 5.

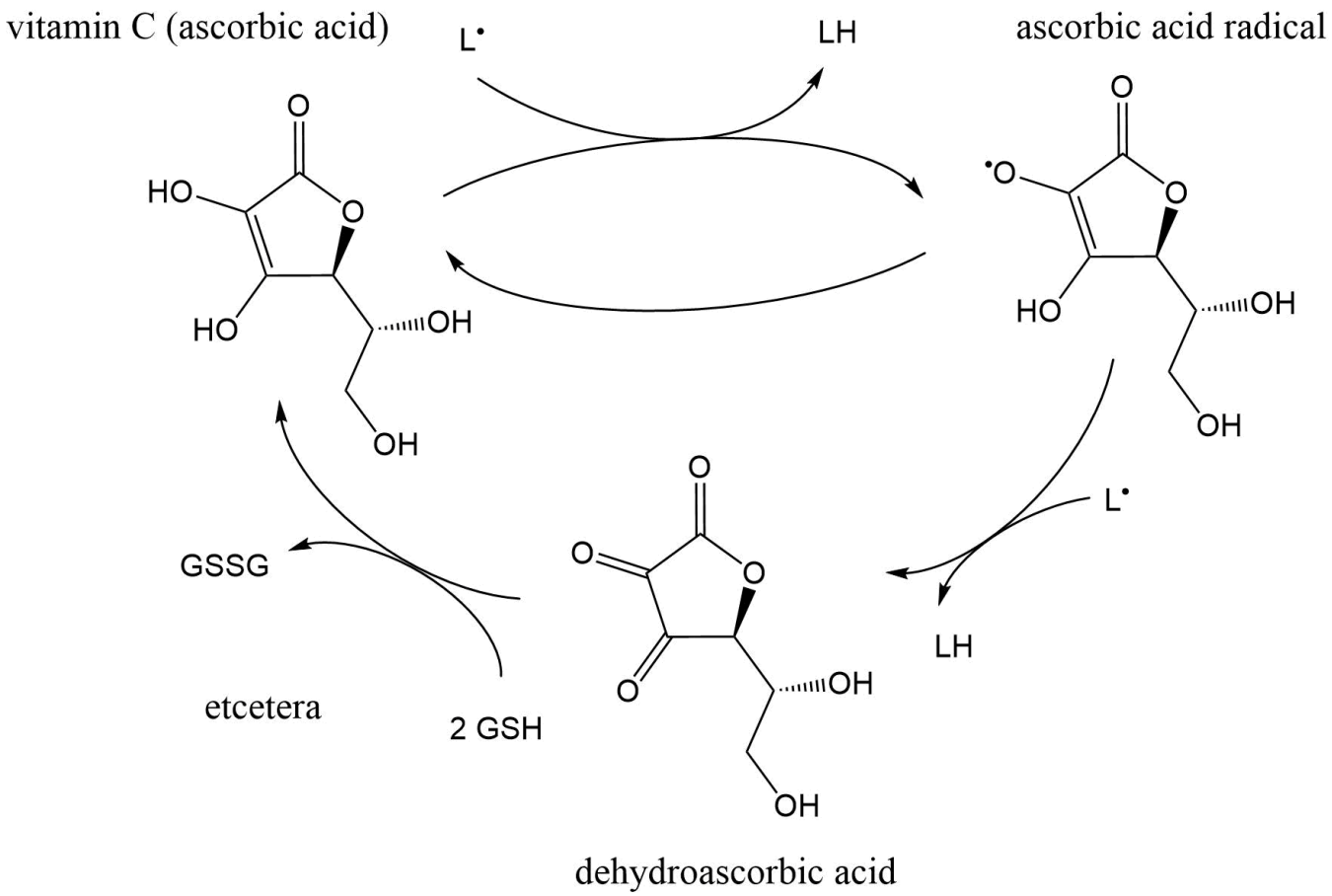


Figure adapted from Niesink et al. (1996) by Steven Droge.

References

Bolton, J.L., Dunlap, T. (2016). Formation and biological targets of quinones: cytotoxic versus cytoprotective effects. Chemical Research in Toxicology 30, 13-37.

Das, K., Roychoudhury, A. (2014). Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. Frontiers in Environmental Science 2, 53.

Edreva, A. (2005). Generation and scavenging of reactive oxygen species in chloroplasts: a submolecular approach.Agriculture, Ecosystems & Environment 106, 119-133.

Klaassen, C. D. (2013). Casarett & Doull's Toxicology: The Basic Science of Poisons, Eighth Edition, McGraw-Hill Professional.

Koppenol, W.H., Traynham, J.G. (1996). Say NO to nitric oxide: nomenclature for nitrogen-and oxygen-containing compounds. In: Methods in Enzymology (Vol. 268, pp. 3-7). Academic Press.

Louise Bolton, J. (2014). Quinone methide bioactivation pathway: contribution to toxicity and/or cytoprotection?. Current Organic Chemistry 18, 61-69.

Nimse, S.B., Pal, D. (2015). Free radicals, natural antioxidants, and their reaction mechanisms. Rsc Advances 5, 27986-28006.

Onyango, A.N. (2016). Endogenous generation of singlet oxygen and ozone in human and animal tissues: mechanisms, biological significance, and influence of dietary components. Oxidative medicine and cellular longevity, 2016.

Niesink, R.J.M., De Vries, J., Hollinger, M.A. (1996). *Toxicology: Principles and Applications*. CRC Press.

Smart, R.C., Hodgson, E. (Eds.). (2018). *Molecular and Biochemical Toxicology*. John Wiley & Sons.

4.2.3.I. Question 1

What are the chances of hydroxyl radicals being formed inside the cell? On what factors does such formation depend?

4.2.3.I. Question 2

Given:

Two oxygen species:

I atmospheric oxygen (O_2)

II singlet oxygen ($^1O_2$)

Which oxygen species contains one or more unpaired electrons, and therefore has radical properties?

I and II

only I

only II

neither I nor II

4.2.3.I. Question 3

Which of the following radicals are detoxified by a-tocopherol (vitamine E)?

I hydroxyl radical, $^{\bullet}OH$

II superoxide anion radical, $O_2^{-\bullet}$

III lipid radical ($L^{\bullet}$)

IV lipid peroxyl radical ($LOO^{\bullet}$)

I and II

III and IV

I, II and III

II, III and IV

4.2.3.I. Question 4

Given:

Three enzymes:

I catalase

II peroxidase

III superoxide dismutase

What enzyme removes hydrogen peroxide?

I and II

I and III

II and III

only III

4.2.3. Oxidative stress - II.

Induction by chemical exposure and possible effects

Author: Frank van Belleghem

Reviewers: Raymond Niesink, Kees van Gestel, Éva Hideg

Learning objectives:

You should be able to

- explain how xenobiotic compounds can lead to an increased production of.
- explain what oxidative stress does with
  - proteins,
  - lipids,
  - DNA and
  - gene regulation.

Keywords: prooxidant-antioxidant balance, bioactivation, oxidative damage,

How xenobiotic compounds induce generation of ROS

The formation of reactive oxygen species (ROS; see section on Oxidative stress I) may involve endogenous substances and chemical-physiological processes as well as xenobiotics. Experimental evidence has shown that oxidative stress can be considered as one of the key mechanisms contributing to the cellular damage of many toxicants. ***Oxidative stress*** has been defined as "a disturbance in the prooxidant-antioxidant balance in favour of the former", leading to potential damage. It is the point at which the production of ROS exceeds the capacity of antioxidants to prevent damage (Klaassen et al., 2013).

Xenobiotics involved in the formation of the superoxide anion radical are mainly substances that can be taken up in so reactive oxygen species -called ***redox cycles***. These include quinones and hydroquinones in particular. In the case of quinones the redox cycle starts with a one-electron reduction step, just as in the case of benzoquinone (Figure 1). The resulting benzosemiquinone subsequently passes the electron received on to molecular oxygen. The reduction of quinones is catalyzed by the NADPH-dependent cytochrome P-450 reductase.

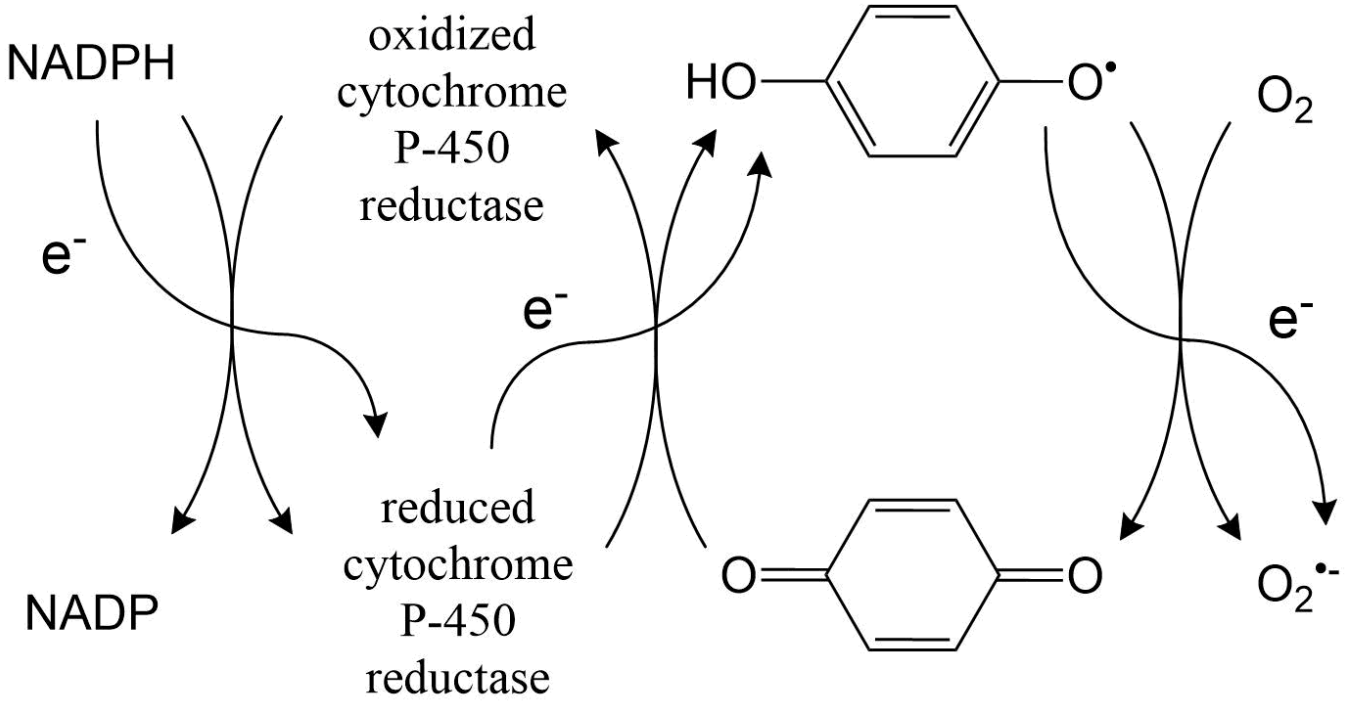


Figure adapted from Niesink et al. (1996) by Steven Droge.

Obviously, hydroquinones can enter a redox cycle via an oxidative step. This step may be catalyzed by enzymes, for example prostaglandin synthase.

Other types of xenobiotic that can be taken up in a redox cycle, are the bipyridyl derivatives. A well-known example is the herbicide paraquat, which causes injury to lung tissue in humans and animals. Figure 2 schematically shows its bioactivation. Other

compounds that are taken up in a redox cycle are nitroaromatics, azo compounds, aromatic hydroxylamines and certain metal (particularly Cu and Zn) chelates.

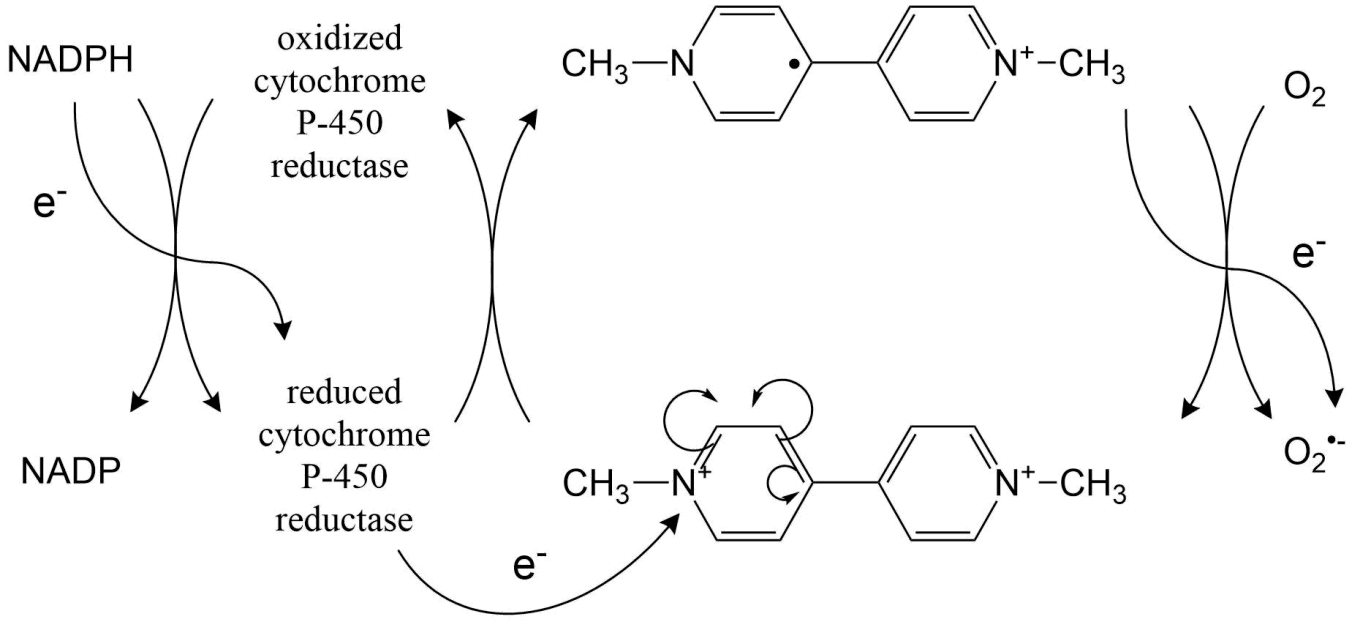


Figure adapted from Niesink et al. (1996) by Steven Droge.

Xenobiotics can enhance ROS production if they are able to enter mitochondria, microsomes, or chloroplasts and ***interact with the electron transport chains***, thus blocking the normal electron flow. As a consequence, and especially if the compounds are electron acceptors, they divert the normal electron flow and increase the production of ROS. A typical example is the cytostatic drug doxorubicin, a well-known chemotherapeutic agent, which is used in treatment of a wide variety of cancers. Doxorubicin has a high affinity for cardiolipin, an important compound of the inner mitochondrial membrane and therefore accumulates at that subcellular location.

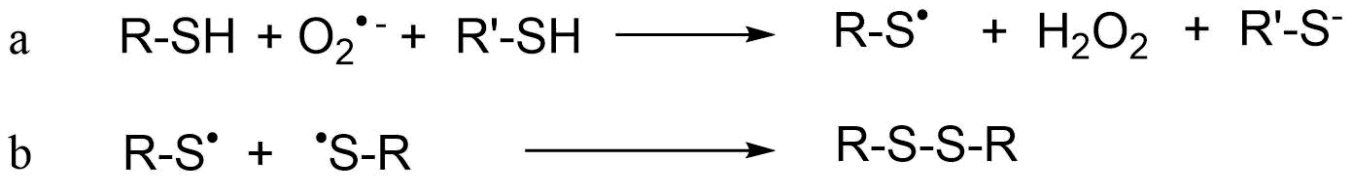
Xenobiotics can cause oxidative damage ***indirectly*** by interfering with the antioxidative mechanisms. For instance it has been suggested that as a non-Fenton metal, cadmium (Cd) is unable to directly induce ROS. However, indirectly, Cd induces oxidative stress by a displacement of redox-active metals, depletion of redox scavengers (glutathione) and inhibition of antioxidant enzymes (protein bound sulfhydryl groups) (Cuypers et al., 2010;Thévenod et al., 2009).

The mechanisms of oxidative stress

As mentioned before, oxidative stress has been defined as "a disturbance in the prooxidant-antioxidant balance in favour of the former". ROS can damage proteins, lipids and DNA via direct oxidation, or through redox sensors that transduce signals, which in turn can activate cell-damaging processes like apoptosis.

Oxidative protein damage

Xenobiotic-induced generation of ROS can damage ***proteins*** through the oxidation of side chains of amino acids residues, the formation of protein-protein cross-links and fragmentation of proteins due to peptide backbone oxidation. The sulfur-containing amino acids cysteine and methionine are particularly susceptible for oxidation. An example of side chain oxidation is the direct interaction of the superoxide anion radical with sulfhydryl (thiol) groups, thereby forming thiyl radicals as intermediates:



As a consequence, glutathione, composed of three amino acids (cysteine, glycine, and glutamate) and an important cellular reducing agent, can be damaged in this way. This means that if the oxidation cannot be compensated or repaired, oxidative stress can lead to depletion of reducing equivalents, which may have detrimental effects on the cell.

Fortunately, antioxidant defence mechanisms limit the oxidative stress and the cell has repair mechanisms to reverse the damage. For example, heat shock proteins (hsp) are able to renature damaged proteins and oxidatively damaged proteins are degraded by the proteasome.

Oxidative lipid damage

Increased concentrations of reactive oxygen radicals can cause membrane damage due to ***lipid peroxidation*** (oxidation of polyunsaturated lipids). This damage may result in altered membrane fluidity, enzyme activity and membrane permeability and transport characteristics. An important feature characterizing lipid peroxidation is the fact that the initial radical-induced damage at a certain site in a membrane lipid is readily amplified and propagated in a ***chain-reaction-like fashion***, thus dispersing the damage across the cellular membrane. Moreover, the products arising from lipid peroxidation (e.g. alkoxy radicals or toxic aldehydes) may be equally reactive as the original ROS themselves and damage cells by additional mechanisms. The chain reaction of lipid peroxidation consists of three steps:

1. Abstraction of a hydrogen atom from a polyunsaturated fatty acid chain by reactive oxygen radicals (*radical formation*, *initiation*).
2. Reaction of the resulting fatty acid radical with molecular oxygen (*oxygenation* or, more specifically, *peroxidation*, *propagation*)
3. These events may be followed by a detoxification process, in which the reaction chain is stopped. This process, which may proceed in several steps, is sometimes referred to as *termination*.

Figure 3 summarizes the various stages in lipid peroxidation.

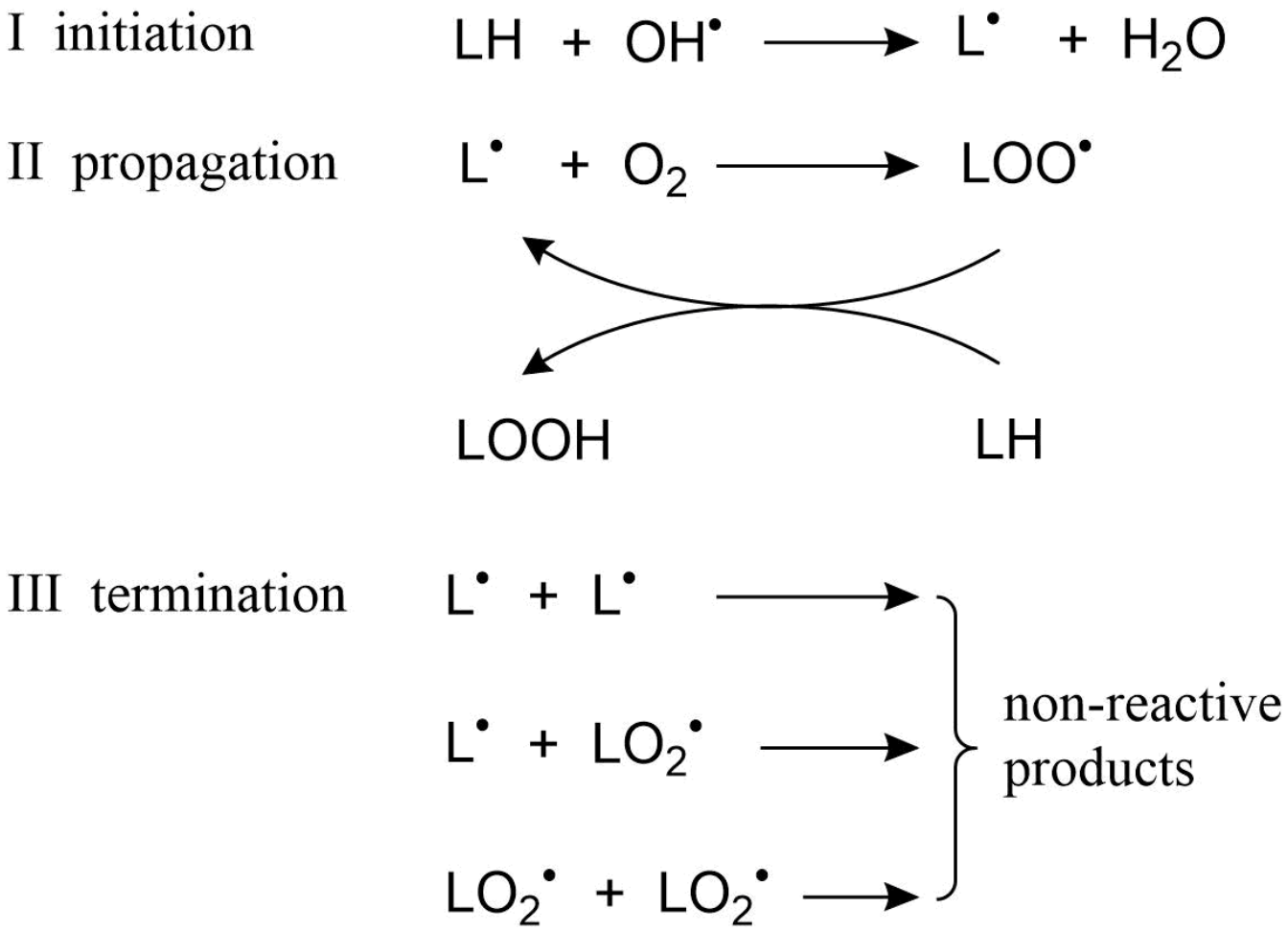


Figure adapted from Niesink et al. (1996) by Steven Droge.

In step II, the peroxidation of biomembranes generates a variety of reactive electrophiles such as epoxides (LOO•) and aldehydes, including malondialdehyde (MDA). MDA is a highly reactive aldehyde which exhibits reactivity toward nucleophiles and can form MDA-MDA dimers. Both MDA and the MDA-MDA dimers are mutagenic and indicative of oxidative damage of lipids from a variety of toxicants.

A classic example of xenobiotic bioactivation to a free radical that initiates lipid peroxidation is the cytochrome P450-dependent conversion of carbon tetrachloride (CCl_4) to generate the trichloromethyl radical ($^{\bullet}CCl_3$) and then the trichloromethyl peroxylradical $CCl_3OO^{\bullet}$. Also the cytotoxicity of free iron is attributed to its function as an electron donor for the Fenton reaction (see section on Oxidative stress I) for instance via the generation of superoxide anion radicals by paraquat redox cycling) leading to the formation of the highly reactive hydroxyl radical, a known initiator of lipid peroxidation.

Oxidative DNA damage

ROS can also oxidize DNA bases and sugars, produce single- or double-stranded DNA breaks, purine, pyrimidine, or deoxyribose modifications and DNA crosslinks. A common modification to DNA is the hydroxylation of DNA bases leading to the formation of oxidized DNA adducts. Although these adducts have been identified in all four DNA bases, guanine is the most susceptible to oxidative damage because it has the lowest oxidation potential of all of the DNA bases. The oxidation of guanine and by hydroxyl radicals leads to the formation 8-hydroxyguanosine (8-OH-dG) (Figure 4).

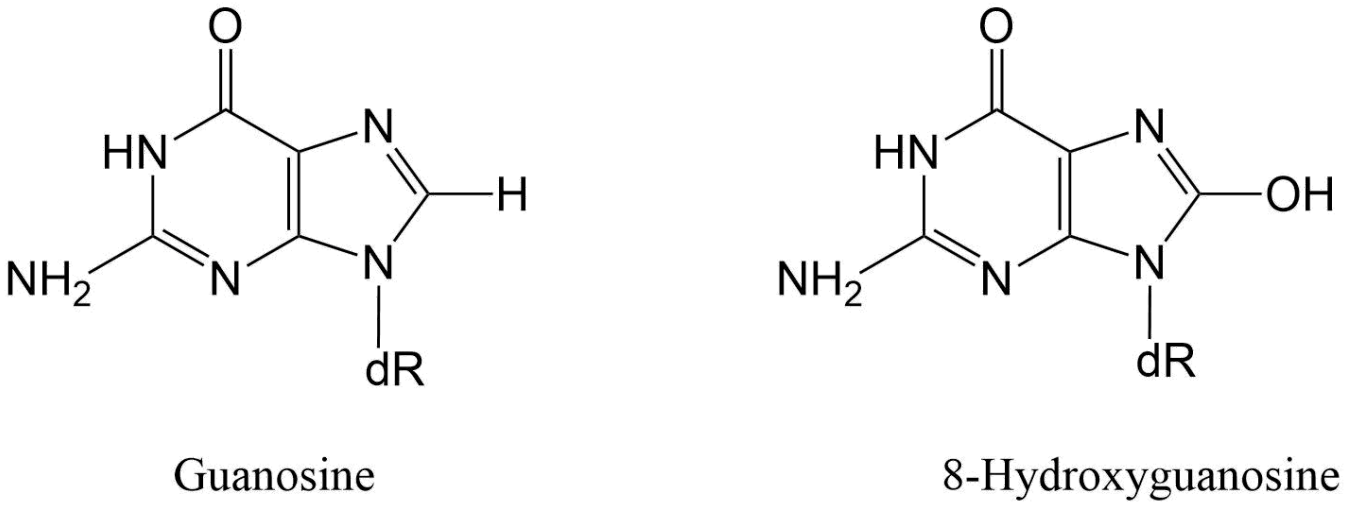


Figure 4. The hydroxylation of guanine. Drawn by Steven Droge.

Oxidation of guanine has a detrimental effect on base paring, because instead of hydrogen bonding with cytosine as guanine normally does, it can form a base pair with adenine. As a result, during DNA replication, DNA polymerase may mistakenly insert an adenosine opposite to an 8-oxo-2'-deoxyguanosine (8-oxo-dG), resulting in a stable change in DNA sequence, a process known as mutagenesis (Figure 5).

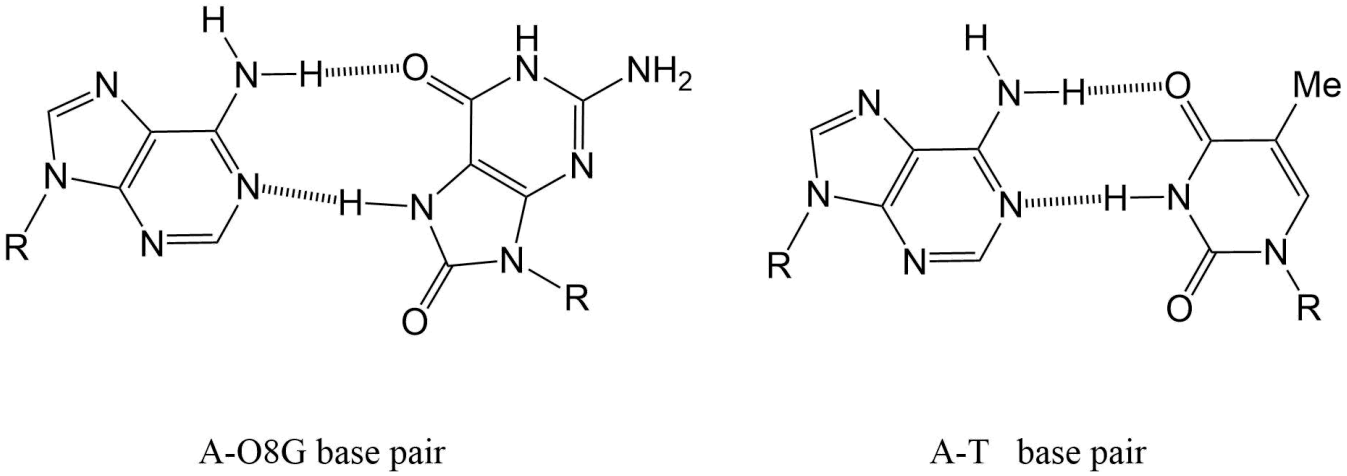


Figure 5. Base paring with 8-oxo-2'-deoxyguanosine (8-oxo-dG). Drawn by Steven Droge.

Fortunately, there is an extensive repair mechanism that keeps mutations to a relatively low level. Nevertheless, persistent DNA damage can result in replication errors, transcription induction or inhibition, induction of signal transduction pathways and genomic instability, events that are possibly involved in carcinogenesis (Figure 6). It has to be mentioned that mitochondrial DNA, is more susceptible to oxidative base damage compared to nuclear DNA due to its proximity to the electron transport chain (a source of ROS), and the fact that mitochondrial DNA is not protected by histones and has a limited DNA repair system.

Figure in preparation

Figure 6. *Oxidative damage by ROS leading to mutations and eventually to tumour formation. Figure adapted from Boelsterli (2002).*

One group of xenobiotics that have clearly been associated with eliciting oxidative DNA damage and cancer are redox-active metals, including Fe(III), Cu(II), Ag(I), Cr(III), Cr(VI), which may entail, as seen before, the production of hydroxyl radicals. Other (non-redox-active) metals that can induce ROS-formation themselves or participate in the reactions leading to endogenously generated ROS are Pb(II), Cd(II), Zn(II), and the metalloid As(III) and As(V). Compounds like polycyclic aromatic hydrocarbons (PAHs), likely the largest family of pollutants with genotoxic effects, require activation by endogenous metabolism to become

reactive and capable of modifying DNA. This activation is brought about by the so-called Phase I biotransformation (see Section on Xenobiotic metabolism and defence).

Genetic detoxifying enzymes, like cytochrome P-450A1, are able to hydrophylate hydrophobic substrates. Whereas this reaction normally facilitates the excretion of the modified substance, some polycyclic aromatic hydrocarbons (PAHs), like benzo[a]pyrene generate semi stable epoxides that can ultimately react with DNA forming mutagenic adducts (see Section on Xenobiotic metabolism and defence). The main regulator of phase I metabolism in vertebrates, the Aryl hydrocarbon receptor (AhR), is a crucial player in this process. Some PAHs, dioxins, and some PCBs (the so-called coplanar congeners; see section on Complex mixtures) bind and activate AhR and increase the activity of phase I enzymes, including cytochrome P-450A1 (CYP1A1), by several fold. This increased oxidative metabolism enhances the toxic effects of the substances leading to increased DNA damage and inflammation (Figure 7).

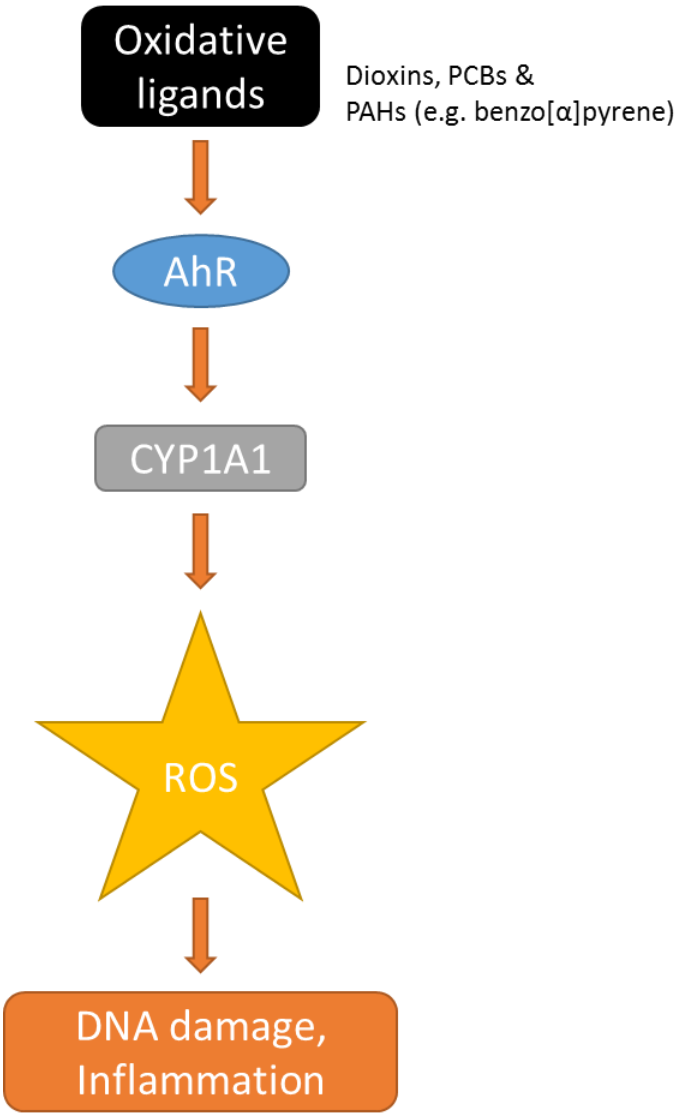


Figure 7. Environmental pollutants such as Dioxines, PCBs, PAHs (such as benzo[a]pyrene) bind to AhR and induce ROS production, DNA damage, and inflammatory cytokine production. Drawn by Frank van Belleghem.

Oxidative effects on cell growth regulation

ROS production and oxidative stress can act both on cell proliferation and apoptosis. It has been demonstrated that low levels of ROS influence signal transduction pathways and alter gene expression.

Figure in preparation

Figure 8. *Role of ROS in altered gene expression. Figure adapted from Klaassen (2013).*

Many xenobiotics, by increasing cellular levels of oxidants, alter gene expression through activation of signaling pathways including cAMP-mediated cascades, calcium-calmodulin pathways, transcription factors such as AP-1 and NF-κB, as well as signaling through mitogen activated protein (MAP) kinases (Figure 8). Activation of these signaling cascades ultimately leads to altered gene expression or a number of genes including those affecting proliferation, differentiation, and apoptosis.

References

Boelsterli, U.A. (2002). *Mechanistic toxicology: the molecular basis of how chemicals disrupt biological targets*. CRC Press.

Cuypers, A., Plusquin, M., Remans, T., Jozefczak, M., Keunen, E., Gielen, H., ... , Nawrot, T. (2010). Cadmium stress: an oxidative challenge. Biometals 23, 927-940.

Furue, M., Takahara, M., Nakahara, T., Uchi, H. (2014). Role of AhR/ARNT system in skin homeostasis. Archives of Dermatological Research 306, 769-779.

Klaassen, C.D. (2013). Casarett & Doull's Toxicology: The Basic Science of Poisons, Eighth Edition, McGraw-Hill Professional.

Niesink, R.J.M., De Vries, J. & Hollinger, M. A. (1996). Toxicology: Principles and Applications. CRC Press.

Thévenod, F. (2009). Cadmium and cellular signaling cascades: to be or not to be? Toxicology and Applied Pharmacology 238, 221-239.

4.2.3.II. Question 1

The herbicide paraquat induces oxidative stress due to

- Interaction with the electron transport chain.

Its involvement in the redox cycle.

It interacts with glutathione.

Its involvement in the Fenton reaction.

4.2.3.II. Question 2

Which biopolymers can undergo damage from reactive oxygen species?

only DNA and proteins

only DNA and membranes

only proteins and membranes

DNA, proteins and membranes

4.2.3.II. Question 3

Given: The three steps of lipid peroxidation:

I initiation

II propagation

III termination

Question: In which step(s) is O_2 involved as a reagent or as a product?

Only I

Only II

I and II

II and III

4.2.3.II. Question 4

Mitochondrial DNA, compared to nuclear DNA, is relatively susceptible to oxidative base damage.

Which of the given alternatives is *not* correct?

The increased susceptibility of mitochondrial DNA is due to:

The proximity of mitochondrial DNA to the electron transport chain

Mitochondrial DNA is not protected by histones

The limited levels of antioxidative compounds inside mitochondria

The limited mitochondrial DNA repair system

4.2.4. Cytotoxicity: xenobiotic compounds causing cell death

Authors: Frank Van Belleghem, Karen Smeets

Reviewers: Timo Hamers, Bas J. Blaauboer

Learning objectives:

You should be able to:

- name the main factors that cause cell death,
- describe the process of necrosis and apoptosis,
- describe the morphological differences between apoptosis and necrosis,
- explain what form of cell death is caused by chemical substances.

Keywords: cell death, apoptosis, necrosis, caspase activation, mitochondrial permeability transition

Description

Cytotoxicity or cell toxicity is the result of chemical-induced macromolecular damage (see the section on Protein inactivation) or receptor-mediated disturbances (see the section on Receptor interactions). Initial events such as covalent binding to DNA or proteins; loss of calcium control or oxidative stress (see the sections on Oxidative stress I and II) can compromise key cellular functions or trigger cell death. Cell death is the ultimate endpoint of lethal cell injury; and can be caused by chemical compounds, mediator cells (*i.e.* natural killer cells) or physical/environmental conditions (*i.e.* radiation, pressure, etc.). The multistep process of cell death involves several regulated processes and checkpoints to be passed before the cell eventually reaches a point of no return, leading to either programmed cell death or apoptosis, or to a more accidental form of cell death, called necrosis. This section describes the cytotoxic process itself, *in vitro* cytotoxicity testing is dealt with in the section on Human toxicity testing - II. In vitro tests.

Chemical toxicity leading to cell death

Cells can actively maintain the intracellular environment within a narrow range of physiological parameters despite changes in the conditions of the surrounding environment. This internal steady-state is termed cellular homeostasis. Exposure to toxic compounds can compromise homeostasis and lead to injury. Cell injury may be direct (primary) when a toxic substance interacts with one or more target molecules of the cell (*e.g.* damage to enzymes of the electron transport chain), or indirect (secondary) when a toxic substance disturbs the microenvironment of the cell (*e.g.* decreased supply of oxygen or nutrients). The injury is called reversible when cells can undergo repair of adaptation to achieve a new viable steady state. When the injury persists or becomes too severe, it becomes irreversible and the cell eventually perishes, thereby terminating cellular functions like respiration, metabolism, growth and proliferation, resulting in cell death (Niesink et al., 1996).

The main factors determining the occurrence of cell death are:

- the nature and concentration of the active toxic compound - in some cases a reactive intermediate - and the availability of that agent at the site of the target molecules;

- the role of the target molecules in the functioning of the cell and/or maintaining the microenvironment;
- the effectiveness of the cellular defence mechanisms in the detoxication and elimination of active agents, in repairing (primary) damage, and in the ability to induce proteins that either promote or inhibit the cell death process.

It is important to realize that also "harmless" substances such as glucose or salt may lead to cell injury and cell death by disrupting the osmotic homeostasis at sufficient concentrations. Even an essential molecule such as oxygen causes cell injury at sufficiently high partial pressures (see the sections on Oxidative stress I and II). Apart from that, all chemicals exert "baseline toxicity" (also called "narcosis") as described in the textbox "narcosis and membrane damage" in the section on Toxicodynamics & Molecular Interactions.

The main types of cell death: necrosis and apoptosis

The two most important types of cell death are necrosis or accidental cell death (ACD) and apoptosis, a form of programmed cell death (PCD) or cell suicide. Cellular imbalances that initiate or promote cell death alone or in combination are oxidative stress, mitochondrial injury or disturbed calcium fluxes. These alterations are reversible at first, but after progressive injury, result in irreversible cell death. Cell death can also be initiated via receptor-mediated signal transduction processes. Apoptotic and necrotic cells differ in both the morphological appearance as well as biochemical characteristics. Necrosis is associated with cell swelling and a rapid loss of membrane integrity. Apoptotic cells shrink into small apoptotic bodies. Leaking cells during necrosis induce inflammatory responses, although inflammation is not entirely excluded during the apoptotic process (Rock & Kono, 2008).

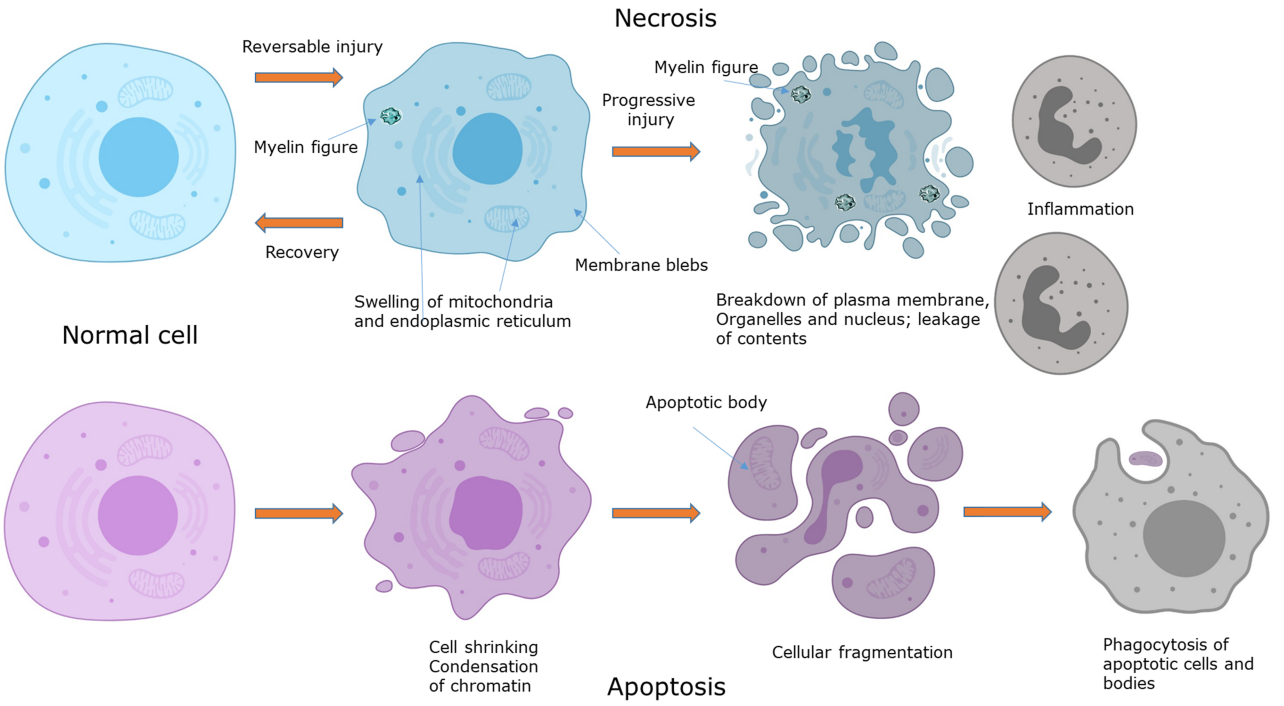
Necrosis

Necrosis has been termed accidental cell death because it is a pathological response to cellular injury after exposure to severe physical, chemical, or mechanical stressors. Necrosis is an energy-independent process that corresponds with damage to cell membranes and subsequent loss of ion homeostasis (in particular Ca^{2+}). Essentially, the loss of cell membrane integrity allows enzymes to leak out of the lysosomal membranes, destroying the cell from the inside. Necrosis is characterized by swelling of cytoplasm and organelles, rupture of the plasma membrane and chromatin condensation (see Figure 1). These morphological appearances are associated with ATP depletion, defects in protein synthesis, cytoskeletal damage and DNA-damage. Besides, cell organelles and cellular debris leak via the damaged membranes into the extracellular space, leading to activation of the immune system and inflammation (Kumar et al., 2015). In contrast to apoptosis, the fragmentation of DNA is a late event. In a subsequent stage, injury is propagated across the neighbouring tissues via the release of proteolytic and lipolytic enzymes resulting in larger areas of necrotic tissue. Although necrosis is traditionally considered as an uncontrolled form of cell death, emerging evidence points out that the process can also occur in a regulated and genetically controlled manner, termed regulated necrosis (Berghe et al., 2014). Moreover, it can also be an autolytic process of cell disintegration after the apoptotic program is completed in the absence of scavengers (phagocytes), termed post-apoptotic or secondary necrosis (Silva, 2010).

Apoptosis

Apoptosis is a regulated (programmed) physiological process whereby superfluous or potentially harmful cells (for example infected or pre-cancerous cells) are removed in a tightly controlled manner. It is an important process in embryonic development, the immune system and in fact, all living tissues. Apoptotic cells shrink and break into small fragments that are phagocytosed by adjacent cells or macrophages without producing an inflammatory response (Figure 3). It can be seen as a form of cellular suicide because cell death is the result of induction of active processes within the cell itself. Apoptosis is an energy-dependent process (it requires ATP) that involves the activation of caspases (cysteine-aspartyl proteases), pro-apoptotic proteins present as zymogens (*i.e.* inactive enzyme precursors that are activated by hydrolysis). Once activated, they function as cysteine proteases and activate other caspases. Caspases can be distinguished into two groups, the initiator caspases, which start the process, and the effector caspases, which specifically lyse molecules that are essential for cell survival (Blanco & Blanco 2017). Apoptosis can be triggered by stimuli coming from within the cell (intrinsic pathway) or from the extracellular medium (extrinsic pathway) as shown in Figure 2. The extrinsic pathway activates apoptosis in response to external stimuli, namely by extracellular ligands binding to cell-surface death receptors (Tumour Necrosis Factor Receptor ((TNFR)), leading to the formation of the death-inducing signalling complex (DISC) and the caspase cascade leading to apoptosis. The intrinsic pathway is activated by cell stressors such as DNA damage, lack of growth factors, endoplasmic reticulum (ER) stress, reactive oxygen species (ROS) burden, replication stress, microtubular alterations and mitotic defects (Galluzzi et al., 2018). These cellular events cause the release of cytochrome c and other pro-apoptotic proteins from the mitochondria into the cytosol via the mitochondrial permeability transition (MPT) pore. This is a

megachannel in the inner membrane of the mitochondria composed of several protein complexes that facilitate the release of death proteins such as cytochrome c. The opening is triggered and tightly regulated by anti-apoptotic proteins, such as B-cell lymphoma-2 (Bcl-2) and pro-apoptotic proteins, such as Bax (Bcl-2 associated X protein) and Bak (Bcl-2 antagonist killer). The intrinsic and extrinsic pathways are regulated by the apoptosis inhibitor protein (AIP) which directly interacts with caspases and suppresses apoptosis. The release of the death protein cytochrome c induces the formation of a large protein structure formed in the process of apoptosis (the apoptosome complex) activating the caspase cascade leading to apoptosis. Other pro-apoptotic proteins oppose to Bcl (SMAC/Diablo) and stimulate caspase activity by interfering with AIP (HtrA2/Omi). HtrA2/Omi also activates caspases and endonuclease G (responsible for DNA degradation, chromatin condensation, and DNA fragmentation). The apoptosis-inducing factor (AIF) is involved in chromatin condensation and DNA fragmentation. Many xenobiotics interfere with the MPT pore and the fate of a cell depends on the balance between pro- and anti-apoptotic agents (Blanco & Blanco, 2017).



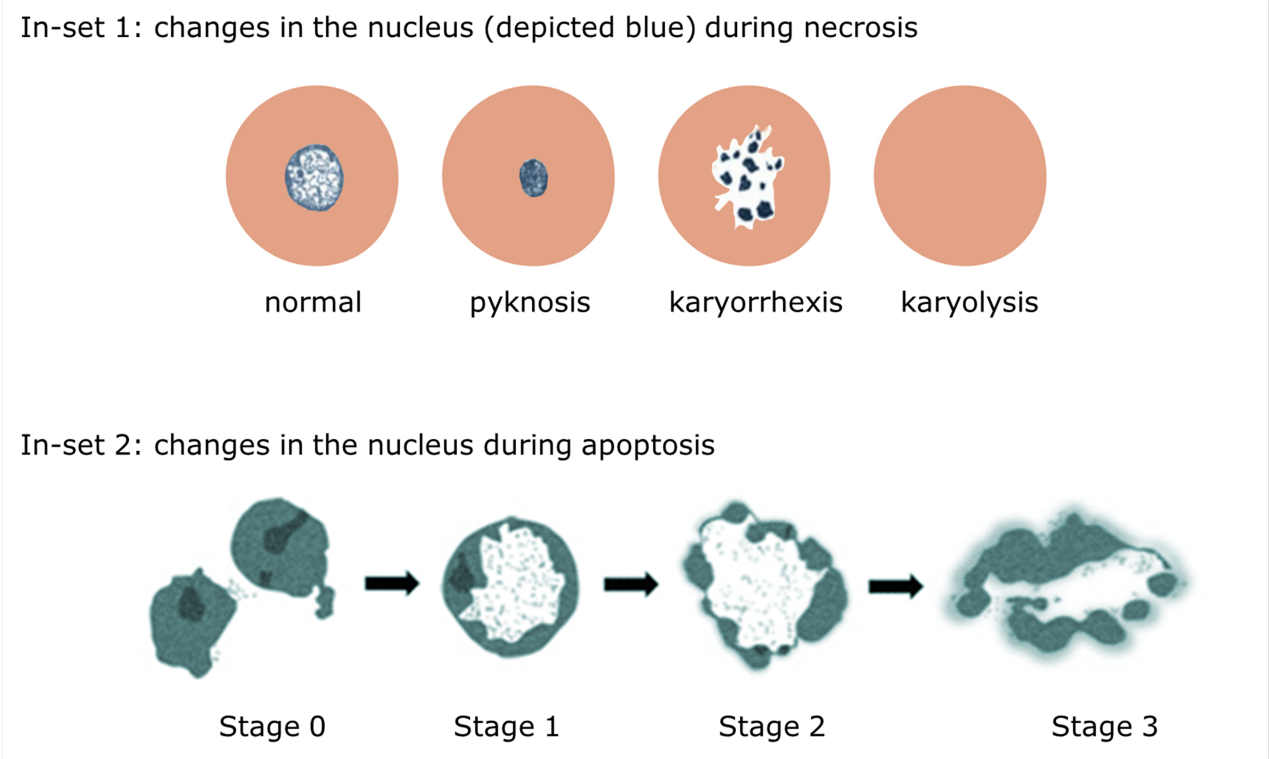


Figure 1. This diagram shows the observable differences between necrotic and apoptotic cell death. Reversible injury is characterized by cytoplasmic enlargement (oncosis), membrane blebbing, swelling of endoplasmic reticula and mitochondria, and the presence of myelin figures (twirled phospholipid masses from damaged cell membranes). Progressive injury leads to the necrotic breakdown of membranes, organelles and the nucleus. The nucleus can thereby undergo shrinking (pyknosis), fragmentation (karyorrhexis) or complete dissolution with loss of chromatin (karyolysis) (see in-set 1). The cell is eventually disrupted, releasing its contents and inducing an inflammatory reaction. In contrast, a cell undergoing apoptosis displays cell

shrinkage, membrane blebbing, and (ring-shaped) chromatin condensation (see in-set 2, image adapted from Toné et al., 2007). The nucleus and cytoplasm break up into fragments called apoptotic bodies, which are phagocytosed by surrounding cells or macrophages.

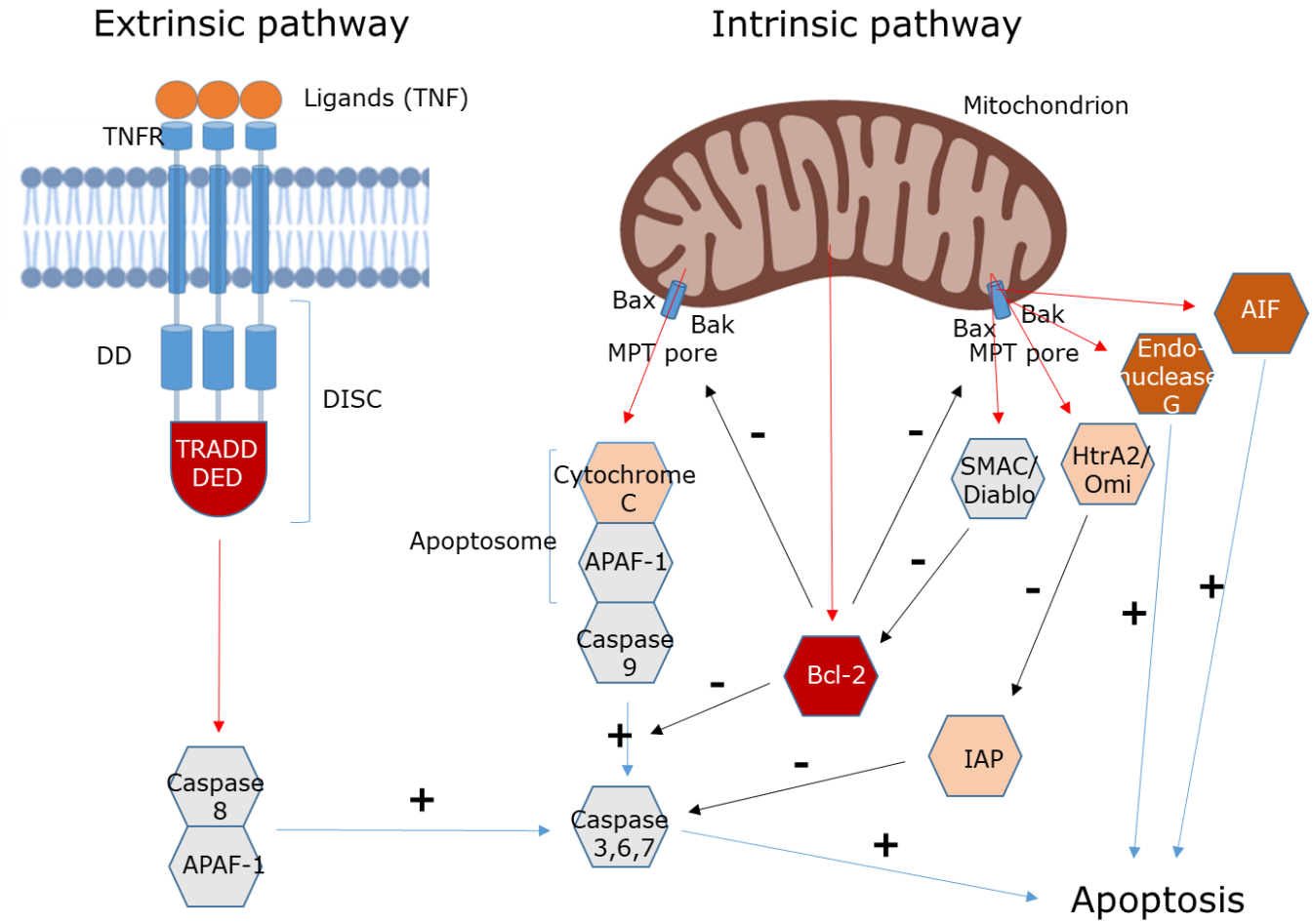


Figure 2. Scheme of the factors involved in the apoptotic process. APAF-1, Apoptosis protease activator factor-1; Bak, Bcl-associated antagonist killer; Bax, Bcl-associated X protein; Bcl-2, B cell lymphoma-2; DD, death domain; DED, death effector domain; IAP, inhibitor apoptosis protein; AIF, apoptosis-inducing factor; TNFR, tumour necrosis factor receptor; TRADD, TNFR-associated death domain. Image adapted from Blanco & Blanco 2017.

What determines the form of cell death caused by chemical substances?

Traditionally, toxic cell death was considered to be uniquely of the necrotic type. The classic example of necrosis is the liver toxicity of carbon tetrachloride (CCl_4) caused by the biotransformation of CCl_4 to the highly reactive radicals (CCl_3• and CCl_3OO•).

Several environmental contaminants including heavy metals (Cd, Cu, CH_3Hg, Pb), organotin compounds and dithiocarbamates can exert their toxicity via induction of apoptosis, likely mediated by disruption of the intracellular Ca^{2+} homeostasis, or induction of mild oxidative stress (Orrenius et al., 2011).

In addition, some cytotoxic substances (*e.g.* arsenic trioxide (As_2O_3)) tend to induce apoptosis at low exposure levels or early after exposure at high levels, whereas they cause necrosis later at high exposure levels. This implicates that the severity of the insult determines the mode of cell death (Klaassen, 2013). In these cases, both apoptosis and necrosis involve the dysfunction of mitochondria, with a central role for the mitochondrial permeability transition (MPT). Normally, the mitochondrial membrane is impermeable to all solutes except for the ones having specific transporters. MPT allows the entry into the mitochondria of solutes with a molecular weight of lower than 1500 Daltons, which is caused by the opening of mitochondrial permeability transition pores (MPTP) in the inner mitochondrial membrane. As these small-molecular-mass solutes equilibrate across the internal mitochondria membrane, the mitochondrial membrane potential ($\Delta\Psi mt$) vanishes (mitochondrial depolarization), leading to uncoupling of

oxidative phosphorylation and subsequent adenosine triphosphate (ATP) depletion. Moreover, since proteins remain within the matrix at high concentration, the increasing colloidal osmotic pressure will result in movement of water into the matrix, which causes swelling of the mitochondria and rupture of the outer membrane. This results in the loss of intermembrane components (like cytochrome c, AIF, HtrA2/Omi, SMAC/Diablo & Endonuclease G) to the cytoplasm. When MPT occurs in a few mitochondria, the affected mitochondria are phagocytosed and the cell survives. When more mitochondria are affected, the release of pro-apoptotic compounds will lead to the caspase activation resulting in apoptosis. When all mitochondria are affected, ATP becomes depleted and the cell will eventually undergo necrosis as shown in Figure 3 (Klaassen et al., 2013).

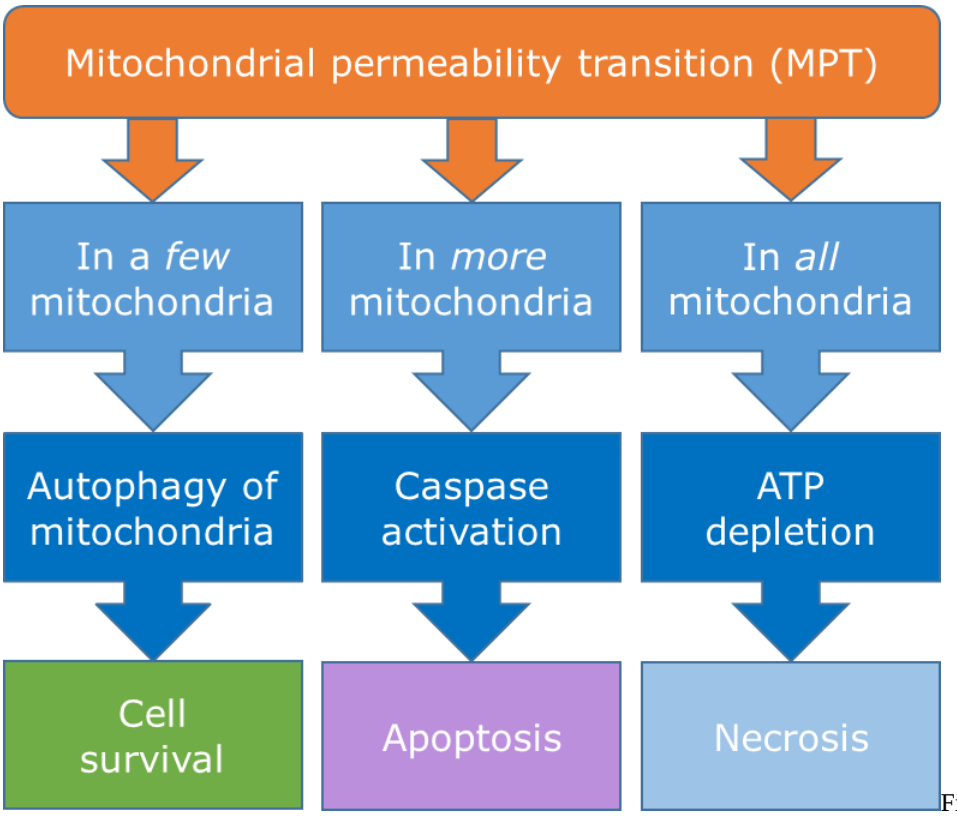


Figure 3. Dose-response relationship of toxicant-induced modes of cell death. The mode of cell death triggered by some toxicants is dose-dependent. Most often, exposure to low doses results in apoptosis, whereas higher levels of the same toxicant might cause necrosis. Image adapted from Klaassen et al., 2013.

References

Berghe, T.V., Linkermann, A., Jouan-Lanhouet, S., Walczak, H., Vandenabeele, P. (2014). Regulated necrosis: the expanding network of non-apoptotic cell death pathways. Nature reviews Molecular Cell Biology 15, 135. https://doi.org/10.1038/nrm3737

Blanco, G., Blanco, A. (2017). Chapter 32 - Apoptosis. *Medical biochemistry*. (pp. 791-796) Academic Press. https://doi.org/10.1016/B978-0-12-803550-4.00032-X

Galluzzi, L., Vitale, I., Aaronson, S.A., Abrams, J.M., Adam, D., Agostinis, P., ... & Annicchiarico-Petruzzelli, M. (2018). Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018. Cell Death & Differentiation, 1. https://doi.org/10.1038/s41418-017-0012-4

Klaassen, C.D., Casarett, L.J., & Doull, J. (2013). *Casarett and Doull's Toxicology: The basic science of poisons* (8th ed.). New York: McGraw-Hill Education / Medical.
ISBN: 978-0-07-176922-8

Kumar, V., Abbas, A.K.,& Aster, J.C. (2015). *Robbins and Cotran pathologic basis of disease, professional edition*. Elsevier Health Sciences. ISBN 978-0-323-26616-1.

Niesink, R.J.M., De Vries, J., Hollinger, M.A. (1996) *Toxicology: Principles and Applications*, (1st ed.). CRC Press. ISBN 0-8493-9232-2.

Orrenius, S., Nicotera, P., Zhivotovsky, B. (2011). Cell death mechanisms and their implications in toxicology. Toxicological Sciences 119, 3-19. https://doi.org/10.1093/toxsci/kfq268

Rock, K.L., Kono, H. (2008). The inflammatory response to cell death. Annu. Rev. Pathmechdis. Mech. Dis. 3, 99-126. https://doi.org/10.1146/annurev.pathmechdis.3.121806.151456

Silva, M.T. (2010). Secondary necrosis: the natural outcome of the complete apoptotic program. FEBS Letters 584, 4491-4499. doi.org/10.1016/j.febslet.2010.10.046

Toné, S., Sugimoto, K., Tanda, K., Suda, T., Uehira, K., Kanouchi, H., ... & Earnshaw, W. C. (2007). Three distinct stages of apoptotic nuclear condensation revealed by time-lapse imaging, biochemical and electron microscopy analysis of cell-free apoptosis. Experimental Cell Research 313, 3635-3644. https://doi.org/10.1016/j.yexcr.2007.06.018

4.2.4. Question 1

Which of the following is a characteristic of necrosis?

- The morphological changes are caused by release of lysosome enzymes.
- It is an energy-dependent process
- It is a programmed response to cellular injury.

It does not lead to inflammation

4.2.4. Question 2

Which of the following is a characteristic of apoptosis?

Membrane bleb formation

Rapid loss of membrane integrity

Swelling of mitochondria

Cell shrinking

4.2.4. Question 3

Which cellular organelles are involved in the initiation of the intrinsic pathway of apoptosis?

Ribosomes

Lysosomes

Mitochondria

Peroxisomes

4.2.4. Question 4

Consider following statements:

- Secondary necrosis is a form of accidental cell death.
- MPT allows the entry of solutes leading to the increase of the volume of the cytoplasm.

Which statement is correct?

- Only I
- Only II
- Neither I, nor II

- Both I and II

4.2.5. Neurotoxicity

Author: Jessica Legradi

Reviewers: Timo Hamers, Ellen Fritsche

Learning objectives

You should be able to

- describe the structure of the nervous system
- explain how neurotransmission works
- mention some modes of action (MoA) by which pesticides and drugs cause neurotoxicity
- understands the relevance of species sensitivity to pesticides
- describe what developmental neurotoxicity (DNT) is

Keywords: Nervous system, Signal transmission, Pesticides, Drugs, Developmental Neurotoxicity

Neurotoxicity

Neurotoxicity is defined as the capability of agents to cause adverse effects on the nervous system. Environmental neurotoxicity describes neurotoxicity caused by exposure to chemicals from the environment and mostly refers to human exposure and human neurotoxicity. Ecological neurotoxicity (eco-neurotoxicity) is defined as neurotoxicity resulting from exposure to environmental chemicals in species other than humans (e.g. fish, birds, invertebrates).

The nervous system

The nervous system consists of the central nervous system (CNS) including the brain and the spinal cord and the peripheral nervous system (PNS). The PNS is divided into the somatic system (voluntary movements), the autonomic (sympathic and parasympathic) system and the enteric (gastrointestinal) system. The CNS and PNS are built from two types of nerve cells, i.e. neurons and glial cells. Neurons are cells that receive, process, and transmit information through electrical and chemical signals. Neurons consist of the soma with the surrounding dendrites and one axon with an axon terminal where the signal is transmitted to another cell (Figure 1A). Compared to neurons, glial cells can have very different appearances (Figure 1B), but are always found in the surrounding tissue of neurons where they provide metabolites, support and protection to neurons without being directly involved in signal transmission.

Figure in preparation

Figure 1. *Structures of a neuron (left; source: https://simple.Wikipedia.org/wiki/Neuron) and of glial cells (right)*

Neurons are connected to each other via synapses. The sending neuron is called the presynaptic neuron whereas the receiving neuron is the postsynaptic neuron. In the synapse, a small space exists between the axon terminal of the presynaptic neuron and a dendrite of the postsynaptic neuron. This space is named synaptic cleft. Both neurons have ion channels that can be opened and closed in the area of the synapse. There are channels selective for chloride, sodium, calcium, potassium, or protons and non-selective channels. The channels can be voltage gated (i.e. they open and close depending on the membrane potential), ligand gated (i.e. they open and close depending on the presence of other molecules binding to the ion channel), or they can be stress activated (i.e. they open and close due to physical stress (stretching)). Ligands that can open or close ion channels are called neurotransmitters. Depending on the ion channel and if it opens or closes upon neurotransmitter binding, a neurotransmitter can inhibit or stimulate membrane depolarization (i.e. inhibitory or excitatory neurotransmitter, respectively). The ligands bind to the ion channel via receptors (link to section on Receptor interaction). Neurotransmitters have very distinct functions and are linked to physical processes like muscle contraction and body heat and to emotional/cognitive processes like anxiety, pleasure, relaxing and learning. The signal transmission via the synapse (i.e. neurotransmission) is illustrated in Figure 2.

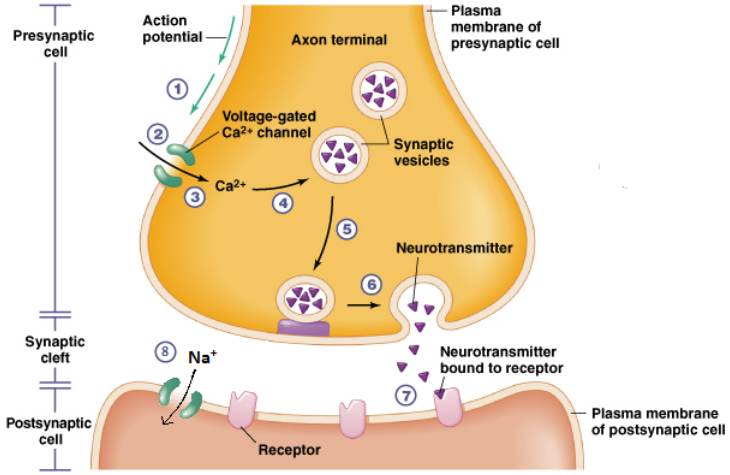


Figure 2: Synaptic neurotransmission by the excitatory neurotransmitter acetylcholinesterase (ACh): 1. action potential arrives at presynaptic neuron; 2. stimulates opening of voltage-gated channels for Ca^{2+}; 3. Ca^{2+}diffuses into the cytoplasm of the presynaptic cell; 4+5. Ca^{2+} causes vesicles containing ACh to move towards the presynaptic membrane; 6. ACh loaded vesicles **fuse** with the membrane, ACh is released and **diffuses across** the synaptic cleft; 7. ACh **temporarily binds** to receptor proteins on the postsynaptic membrane; causing ligand-gated ion channels for Na^+ to open; 8. Na^+ diffuses through postsynaptic membrane, **depolarizes the membrane** and generates an action potential. Source: http://biology4alevel.blogspot.com/2016/06/122-synapses.html

The cell membrane of a neuron contains channels that allow ions to enter and exit the neuron. This flow of ions is used to send signals from one neuron to the other. The difference in concentration of negatively and positively charged ions on the inner and outer side of the neuronal membrane creates a voltage across the membrane called the membrane potential. When a neuron is at rest (i.e. not signalling), the inside charge of the neuron is negative relative to the outside. The cell membrane is then at its resting potential. When a neuron is signalling, however, changes in ion inflow and outflow of ions lead to a quick depolarization followed by a repolarization of the membrane potential called action potential. A video showing how the action potential is produced can be found here.

Neurons can be damaged via substances that damage the cell body (neuronopathy), the axon (axonopathy), or the myelin sheet or glial cells (myelopathy). Aluminum, arsenic, methanol, methylmercury and lead can cause neuropathy. Acrylamide is known to specifically affect axons and cause axonopathy.

Neurotransmitter system related Modes of Action of neurotoxicity

Some of the modes of action relevant for neurotoxicity are disturbances of electric signal transmission and inhibition of chemical signal transmission, mainly through interference with the neurotransmitters. Pesticides are mostly designed to interfere with neurotransmission.

1. Interfering with Ion channels (see section on Receptor interaction)

Pesticides such as DDT bind to open sodium channels in neurons, which prevents closing of the channels and leads to over-excitation. Pyrethroids, such as permethrin, increase the time of opening of the sodium channels, leading to similar symptoms. Lindane, cyclodiene insecticides like aldrin, dieldrin and endrin ("drins") and phenyl-pyrazols such as fipronil block GABA-mediated chloride channels and prevent hyperpolarization. GABA (gamma-aminobutyric acid) is an inhibitory neurotransmitter which is linked to relaxation and calming. It stimulates opening of chloride channels causing the transmembrane potential to become more negative (i.e. hyperpolarization), thereby increasing the depolarisation threshold for a new action potential. Blockers of GABA-mediated chloride channels prevent the hyperpolarizing effect of GABA, thereby decreasing the inhibitory effect of GABA. Neonicotinoids (e.g., imidacloprid) mimic the action of the excitatory neurotransmitter ACh by activating the nicotinic acetylcholine receptors (nAChR) in the postsynaptic membrane. These compounds are specifically designed for displaying a high affinity to insect nAChR.

Many human drugs, like sedatives also bind to neuro-receptors. Benzodiazepine drugs activate GABA-receptors causing hyperpolarization (activating GABA). Tetrahydrocannabinol (THC), which is the active ingredient in cannabis, activates the

cannabinoid receptors also causing hyperpolarization. Compounds activating the GABA or cannabinoid receptors induce a strong feeling of relaxation. Nicotine binds and activates the AChR, which can help to concentrate.

2. AChE inhibition

Another very common neurotoxic mode of action is the inhibition of acetylcholinesterase (AChE). Organophosphate insecticides like dichlorvos and carbamate insecticides like propoxur bind to AChE, and hence prevent the degradation of acetylcholine in the synaptic cleft, leading to overexcitation of the post-synaptic cell membrane (see also section on Protein interaction).

3. Blocking Neurotransmitter uptake

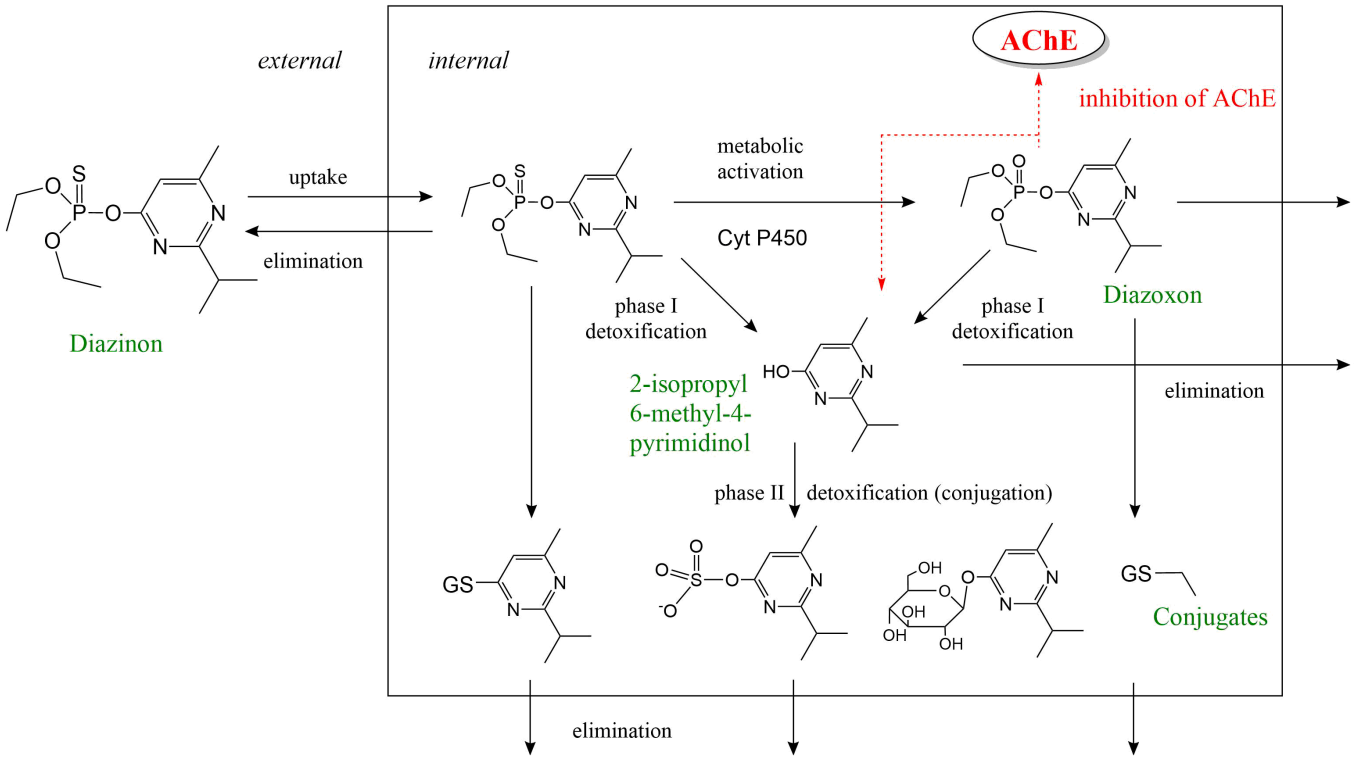
MDMA (3,4-methylenedioxymethamphetamine, also known as ecstasy or XTC) and cocaine block the re-uptake of serotonin, norepinephrine and to a lesser amount dopamine into the pre-synaptic neuron, thereby increasing the amount of these neurotransmitters in the synaptic cleft. Amphetamines also increase the amount of dopamine in the cleft by stimulating the release of dopamine form the vesicles. Dopamine is a neurotransmitter which is involved in pleasure and reward feelings. Serotonin or 5-hydroxytryptamine is a monoamine neurotransmitter linked to feelings of happiness, learning, reward and memory.

Long term exposure

When receptors are continuously activated or when neurotransmitter levels are continuously elevated, the nervous system adapts by becoming less sensitive to the stimulus. This explains why drug addicts have to increase the number of drugs taken to get to the desired state. If no stimulant is taken, withdrawal symptoms occur from the lack of stimulus. In most cases, the nervous system can recover from drug addiction.

Species Sensitivity in Neurotoxicity

Differences in species sensitivity can be explained by differences in metabolic capacities between species. Most compounds need to be bio-activated, i.e. being biotransformed into a metabolite that causes the actual toxic effect. For example, most organophosphate insecticides are thio-phosphoesters that require oxidation prior to causing inhibition of AChE. As detoxification is the dominant pathway in mammals and oxidation is the dominant pathway in invertebrates, organophosphate insecticides are typically more toxic to invertebrates than to vertebrates (see Figure 3). Other factors important for species sensitivity are uptake and depuration rate.



https://bio.libretexts.org/@go/page/108679

Figure 3: *Mechanism of action of an AChE inhibitor on the example of the insecticide diazinon. After oxidation catalyzed by cytochrome p450 monooxygenases, Diazinon is metabolized into diazoxon which can inhibit acetylcholinesterase. Via further phase I and phase II metabolization steps the molecule is eliminated. Drawn by Steven Droge.*

Developmental neurotoxicity

Developmental neurotoxicity (DNT) particularly refers to the effects of toxicants on the developing nervous system of organisms. The developing brain and nervous system are supposed to be more sensitive to toxic effects than the mature brain and nervous system. DNT studies must consider the temporal and regional occurrence of critical developmental processes of the nervous system, and the fact that early life exposure can lead to long-lasting neurotoxic effects or delays in neurological development. Species differences are also relevant for DNT. Here, developmental timing, speed, or cellular specificities might determine toxicity.

4.2.5. Question 1

What are the two major cell types found in the nervous system?

4.2.5. Question 2

What does GABA do?

4.2.5. Question 3

How does AChE inhibition work?

4.2.5. Question 4

What makes invertebrates more sensitive to organophosphate insecticides?

4.2.5. Question 5

Why is DNT important to study?

4.2.6. Effects of herbicides

Author: Nico M. van Straalen

Reviewers: Cornelia Kienle, Henk Schat

Learning objectives

You should be able to

- Explain the different ways in which herbicides are applied in modern agriculture
- Enumerate the eight major modes of action of herbicides
- Provide some examples of side-effects of herbicides

Keywords: Amino acid inhibitor, growth regulator, photosynthesis inhibitor, pre-emergence application, selectivity

Introduction

Herbicides are pesticides (see section on Crop protection products) that aim to kill unwanted weeds in agricultural systems, and weeds growing on infrastructure such as pavement and train tracks. Herbicides are also applied to the crop itself, e.g. as a pre-harvest treatment in crops like potato and oilseed rape, to prevent growth of pathogens on older plants, or to ease mechanical harvest. In a similar fashion, herbicides are used to destroy grass of pastures in preparation of their conversion to cropland. These applications are designated "desiccation". Finally herbicides are used to kill broad-leaved weeds in pure grass-fields (e.g. golf courts).

Herbicides represent the largest volume of pesticides applied to date (about 60%), partly because mechanical and hand-executed weed control has declined considerably. The tendency to limit soil tillage (as a strategy to maintain a diverse and healthy soil life) has also stimulated the use of chemical herbicides.

Herbicides are obviously designed to kill plants and therefore act upon biochemical targets that are specific to plants. As the crop itself is also a plant, selectivity is a very important issue in herbicide application. This is achieved in several ways.

- Application of herbicides before emergence of the crop (pre-emergence application). This will keep the field free of weeds before germination, while the closed canopy of the crop prevents the later growth of weeds. This strategy is often applied in fast-growing crops that make a high canopy with a lot of shading on ground level, such as maize. Examples of herbicides used in pre-emergence application are glyphosate and metolachlor. Also the selectivity of seedling growth inhibitors such as EPTC is due to the fact that these compounds are applied as part of a soil preparation and act on germinating plants before the crop emerges.
- Broad-leaved plants are more susceptible to herbicides that rely on contact with the leaves, because they intercept more of a herbicide spray than small-leaved plants such as grass. This type of selectivity allows some herbicides to be used in grassland and cereal crops, to control broad-leaved weeds; the herbicide itself is not intercepted by the crop. Examples are the chlorophenoxy-acetic acids such as MCPA and 2,4-D.
- In some cases the crop plant is naturally tolerant to a herbicide due to specific metabolic pathways. The selectivity of ACCase inhibitors such as diclofop-methyl, fenoxaprop-ethyl, and fluazifop-butyl is mostly due to this mechanism. These compounds inhibit acetyl-CoA carboxylases, a group of enzymes essential to fatty acid synthesis. However, in wheat the herbicidal compounds are quickly hydrolysed to non-toxic metabolites, while weeds are not capable of such detoxification. This allows such herbicides to be used in wheat fields. Another type of physiological selectivity is due to differential translocation, that is, some plants quickly transport the herbicide throughout the plant, enabling it to exert toxicity in the leaves, while others keep the substance in the roots and so remain less susceptible.
- Several crops have been genetically modified (gm) to become resistant to herbicides; one of the best-known modifications is the insertion of an altered version of the enzyme EPSP synthase. This enzyme is part of the shikimate pathway and is specifically inhibited by glyphosate (Figure 1). The modified version of the enzyme renders the plant insensitive to glyphosate, allowing herbicide use without damage to the crop. Various plant species have been modified in this way, although their culture is limited to countries that allow gm-crops (USA and many other countries, but not European countries).

Classification by mode of action

The diversity of chemical compounds that have been synthesized to attack specific biochemical targets in plants is enormous. In an attempt to classify herbicides by mode of action a system of 22 different categories is often used (Sherwani et al. 2015). Here we present a simplified classification specifying only eight categories (Plant & Soil Sciences eLibrary 2019, Table 1).

Table 1. Classification of herbicides by mode of action

No.	Class (mode of action)	Examples of chemical groups	Example of active ingredient
1	Amino acid synthesis inhibitors	Sulfonylureas, imidazolones, triazolopyrimidines, epsp synthase inhibitors	Glyphosate
2	Seedling growth inhibitors	Carbamothiates, acetamides, dinitroanilines	EPTC
3	Growth regulators (interfere with plant hormones)	Phenoxy-acetic acids, benzoic acid, carboxylic acids, picolinic acids	2,4-D
4	Inhibitors of photosynthesis	Triazines, uracils, phenylureas, benzothiadiazoles, nitriles, pyridazines	Atrazine
5	Lipid synthesis inhibitors	Aryloxyphenoxypropionates, cyclohexanediones	Sethoxydim
6	Cell membrane disrupters	Diphenylethers, aryl triazolinones, phenylphthalamides, bipyridilium	Paraquat
7	Inhibitors of protective pigments	Isoxazolidones, isoxazoles, pyridazinones	Mesotrione
8	Unknown	Chemical compounds with proven herbicide efficacy but unknown mode of action	Ethofumesate

To illustrate the diversity of herbicidal mode of action, two examples of well-investigated mechanisms are highlighted here.

Plants synthesize aromatic amino acids using the shikimate pathway. Also bacteria and fungi avail of this pathway, but it is not present in animals. They must obtain aromatic amino acids through their diet. The first step in this pathway is the conversion of shikimate-3-phosphate and phosphoenolpyruvate (PEP) to 5-enolpyruvylshikimate-3-phosphate (EPSP), by the enzyme EPSP synthase (Figure 1). EPSP is subsequently dephosphorylated and forms the substrate for the synthesis of aromatic amino acids such as phenylalanine, tyrosine and tryptophan.

Glyphosate bears a structural resemblance to PEP and competes with PEP as a substrate for EPSP synthase. However, in contrast to PEP it binds firmly to the active site of the enzyme and blocks its activity. The ensuing metabolic deficiency quickly leads to loss of growth potential of the plant.

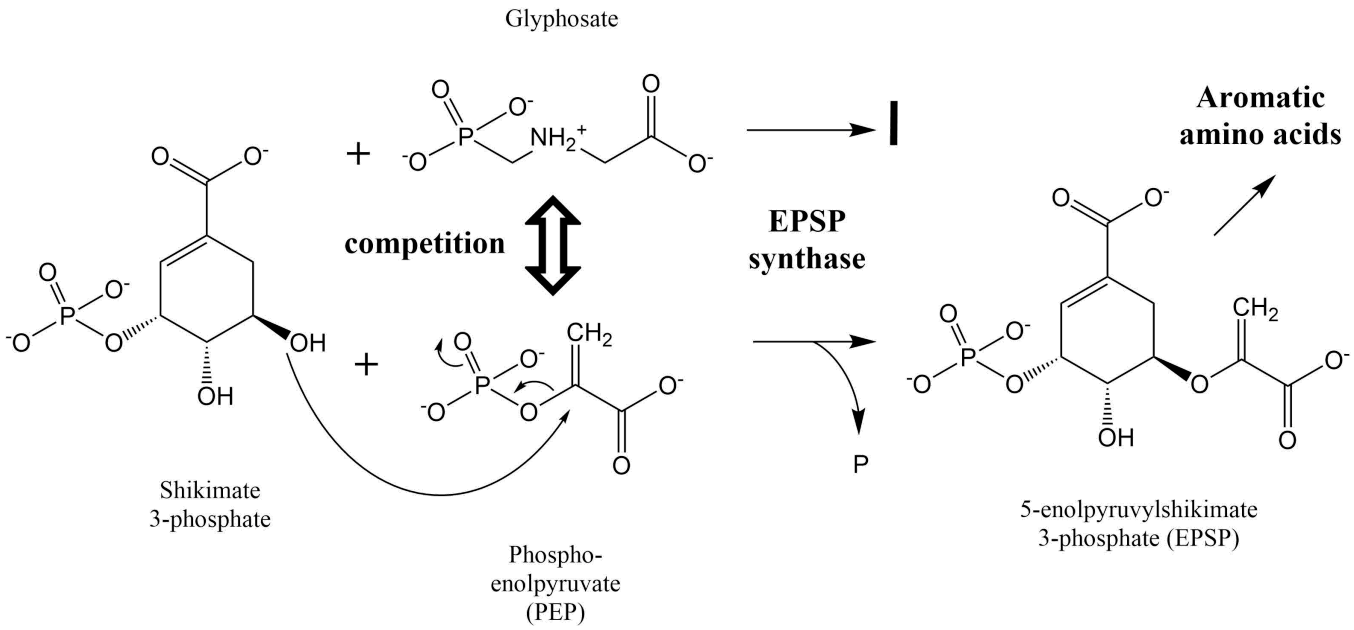


Figure 1. The first step in the shikimate pathway used by plants to synthesize aromatic amino acids. The enzyme EPSP synthase is inhibited by glyphosate due to competitive interaction with PEP. Redrawn by Steven Droge.

Another very well investigated mode of herbicidal action is photosynthesis inhibition by atrazine and other symmetrical triazines. In contrast to glyphosate, atrazine can only act in aboveground plants with active photosynthesis. Sunny weather stimulates the action of such herbicides. The action of atrazine is due to binding to the D1 quinone protein of the electron transport complex of photosystem II sitting in the inner membrane of the chloroplast (see Figure 1 in Giardi and Pace, 2005). Photosystem II (PSII) is a complex of macromolecules with light harvesting and antenna units, chlorophyll P680, and reaction centers that capture light energy and use it to split water, produce oxygen and transfer electrons to photosystem I, which uses them to eventually produce reduction equivalents. The D1 quinone has a "herbicide binding pocket" and binding of atrazine to this site blocks the function of PSII. A single amino acid in the binding pocket is critical for this; alterations in this amino acid provide a relatively easy possibility for the plant to become resistant to triazines.

Side-effects

Most herbicides are polar compounds with good water solubility, which is a crucial property for them to be taken up by plants. This implies that herbicides, especially the more persistent ones, tend to leach to groundwater and surface water and are sometimes also found in drinking water resources. Given the large volumes applied in agriculture, concern has arisen that such compounds, despite them being designed to affect only plants, might harm other, so called "non-target" organisms.

In agricultural systems and their immediate surroundings, complete removal of weeds will reduce plant biodiversity, with secondary effects on plant-feeding insects and insectivorous birds. In the short term however herbicides will increase the amount of dead plant remains on the soil, which may benefit invertebrates that are less susceptible to the herbicidal effect, and find shelter in plant litter and feed on dead organic matter. Studies show that there is often a positive effect of herbicides on Collembola, mites and other surface-active arthropods (e.g. Fratello et al. 1985). Other secondary effects may occur when herbicides reach field-bordering ditches, where suppression of macrophytes and algae can affect populations of macro-invertebrates such as gammarids and snails.

Direct toxicity to non-target organisms is expected from broad-spectrum herbicides that kill plants due to a general mechanism of toxicity. This holds for paraquat, a bipyridilium herbicide (cf. Table 1) that acts as a contact agent and rapidly damages plant leaves by redox-cycling; enhanced by sunshine, it generates oxygen radicals that disrupt biological membranes. Paraquat is obviously toxic to all life and represents an acute hazard to humans. Consequently, its use as a herbicide is forbidden in the EU since 2007.

In other cases the situation is more complex. Glyphosate, the herbicide with by far the largest application volume worldwide is suspect of ecological side-effects and has even been labelled "a probable carcinogen" by the IUCR (Tarazona et al., 2017). However, glyphosate is an active ingredient contained in various herbicide formulations, e.g. Roundup Ready, Roundup 360 plus,

etc. Evidence indicates that most of the toxicity attributed to glyphosate is actually due to adjuvants in the formulation, specifically polyethoxylated tallowamines (Mesnage et al., 2013).

Another case of an unexpected side-effect from a herbicide is due to atrazine. In 2002 a group of American ecologists (Hayes et al., 2002) reported that the incidence of developmental abnormalities in wild frogs was correlated with the volume of atrazine sold in the area where frogs were monitored, across a large number of sites in the U.S. Male *Rana pipiens* exposed to atrazine in concentrations higher than 0.1 µg/L during their larval stages showed an increased rate of feminization, i.e. the development of oocytes in the testis. This would be due to induction of aromatase, a cytochrome P450 activity responsible for the conversion of testosterone to estradiol.

Finally the development of resistance may also be considered an undesirable side-effect. There are currently (2018) 499 unique cases (255 species of plant, combined with 167 active ingredients) of herbicide resistance, indicating the agronomical seriousness of this issue. A full discussion of this topic falls, however, beyond the scope of this module.

Conclusions

Herbicides are currently an indispensable, high-volume component of modern agriculture. They represent a very large number of chemical groups and different modes of action, often plant-specific. While some of the older herbicides (paraquat, atrazine, glyphosate) have raised concern regarding their adverse effects on non-plant targets, the development of new chemicals and the discovery of new biochemical targets in plant-specific metabolic pathways remains an active field of research.

References

Fratello, B. et al. (1985) Effects of atrazine on soil microarthropods in experimental maize fields. Pedobiologia 28: 161-168.

Giardi, M.T., Pace, E. (2005) Photosynthetic proteins for technological applications. Trends in Biotechnology 23, 257-263.

Hayes, T., Haston, K., Tsui, M., Hoang, A., Haeffele, C., Vonk, A. (2002). Feminization of male frogs in the wild. Nature 419, 895-896.

Mesnage, R., Bernay, B., Séralini, G.-E. (2013). Ethoxylated adjuvants of glyphosate-based herbicides are active principles of human cell toxicity. Toxicology 313, 122-128.

Plant & Soil Sciences eLibrary (2019), https://passel.unl.edu.

Sherwani, S.I., Arif, I.A., Khan, H.A. (2015). Modes of action of different classes of herbicides. In: Price, J., Kelton, E., Suranaite, L. (Eds.). Herbicides. Physiological Action and Safety. Chapter 8, IntechOpen.

Tarazona, J.V., Court-Marques, D., Tiramani, M., Reich, H., Pfeil, R., Istace, F., Crivellente, F. (2017). Glyphosate toxicity and carcinogenicity: a review of the scientific basis of the European Union Assessment and its differences with IARC. Archives of Toxicology 91, 2723-2743.

4.2.6. Question 1

Define what is meant by a pre-emergence herbicide and why this is useful in agronomy?

4.2.6. Question 2

With a herbicide application in agriculture you want to kill unwanted plants among a crop that is in itself also a plant. How is this possible?

4.2.6. Question 3

Enumerate the eight major modes of action of herbicides

4.2.6. Question 4

Can herbicides cause adverse effects on non-plant species?

4.2.7. Chemical carcinogenesis and genotoxicity

Author: Timo Hamers

Reviewer: Frederik-Jan van Schooten

Learning objectives

You should be able to

- describe the three different phases in cancer development and understands how compounds can stimulate the corresponding processes in these phases
- explaion the difference between base pair substitutions and frameshift mutations both at DNA and the protein level
- describe the principle of bioactivation, which distinguishes indirect from direct mutagenic substances
- explain the difference between mutagenic and non-mutagenic carcinogens

Key words: Bioactivation; Mutation; Tumour promotion; Tumour progression; Ames test

Chemical carcinogenesis

Cancer is a collective name for multiple diseases sharing a common phenomenon that cell division is out of the control by growth-regulating processes. The consequent, autonomic growing cells are usually concentrated in a neoplasm (often referred to a tumour) but can also be diffusely dispersed, for instance in case of leukaemia or a mesothelioma. Benign tumours refer to neoplasms that are encapsulated and do not distribute through the body, whereas malign tumours cause metastasis , i.e. spreading of carcinogenic cells through the body causing new neoplasms at distant. The term benign sounds more friendly than it actually is: benign tumours can be very damaging to organs which are limited in available space (e.g. the brain in the skull) or to organs that can be obstructed by the tumour (e.g. the gut system).

The process of developing cancer (carcinogenesis) is traditionally divided in three phases, i.e.

1. the **initiation** phase, in the genetic DNA of a cell is permanently changed, resulting in daughter cells that genetically differ from their parent cells;
2. the **promotion** phase, in which the cell loses its differentiation and gains new characteristics causing increased proliferation;
3. the **progression** phase, in which the tumour invades surrounding tissues and causes metastasis.

Chemical carcinogenesis means that a chemical substance is capable of stimulating one or more of these phases. Carcinogenic compounds are often named after the phase that they affect, i.e. initiators (also called mutagens), tumour promotors, and tumour progressors. It is important to realize that many substances and processes naturally occurring in the body can also stimulate the different phases, i.e. inflammation and exposure to sun light may cause mutations, some endogenous hormones can act as very active promotors in hormone-sensitive cancers, and spontaneous mutations may stimulate the tumour progression phase.

Point mutations

Gene mutations (aka point mutations) are permanent changes in the order of the nucleotide base-pairs in the DNA. Based on what happens at the DNA level, point mutations can be divided in three types, i.e. a replacement of an original base-pair by another base-pair (**base-pair substitution**), the **insertion** of an extra base-pair or the **deletion** of an original base-pair (Figure 1). In a coding part of DNA, three adjacent nucleotides on a DNA strand (i.e. a triplet) form a codon that encodes for an amino acid in the ultimate protein. Because insertions and deletions cause a shift in these triplet reading frames with one nucleotide to the left or to the right, respectively, these point mutations are also called **frame-shift mutations**.

Based on what happens at the protein level for which a gene encodes, point mutations can also be divided into three types. A **missense** mutation means that the mutated gene encodes for a different protein than the wildtype gene, a **nonsense** mutation means that the mutation introduces a STOP codon that interrupts gene transcription resulting in a truncated protein, and a **silent** mutation means that the mutated gene still encodes for exactly the same protein, despite the fact that the genetic code has been changed. Silent mutations are always base-pair substitutions, because the triplet structure of the DNA has not been damaged.

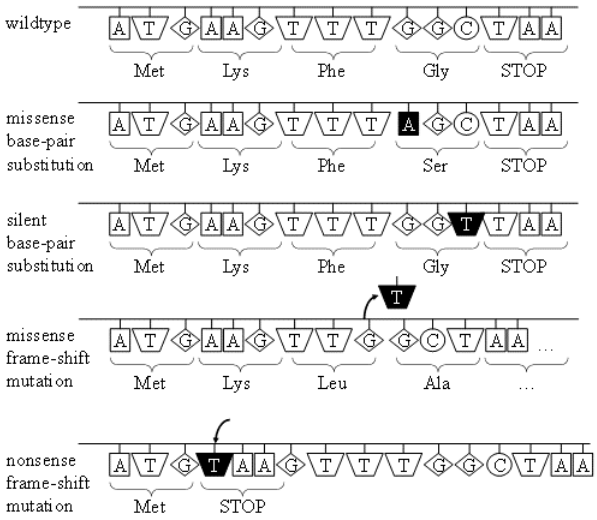


Figure 1: Examples of missense, nonsense en silent mutations at the polypeptide level, based on base-pair substitutions en frame-shift mutations at the genomic DNA level.

A very illustrative example of the difference between a base-pair substitution and a frameshift mutation at the level of protein expression is the following "wildtype" sentence, consisting of only three letter words representing the triplets in the genomic DNA:
The fat cat ate the hot dog.
Imagine that the letter t in cat is replaced by an r due to a base-pair substitution. The sentence then reads:
The fat car ate the hot dog.
This sentence clearly has another meaning, i.e. it contains missense information.
Imagine now that the letter a in fat is replaced by an e due to a base-pair substitution. The sentence then reads:
The fet cat ate the hot dog.
This sentence clearly contains a spelling error (i.e. a mutation), but it's meaning has not changed, i.e. it contains a silent mutation.
Imagine now that an additional letter m causes a frameshift in the word fat, due to an insertion. The sentence then reads:
The fma tca tat eth eho tdo.
This sentence clearly has another meaning, i.e. it contains missense information.
Similarly, leaving out the letter a in fat also causes a frameshift mutation, due to a deletion. The sentence then reads:
The ftc ata tet heh otd og.
Again, this sentence clearly has another meaning, i.e. it contains missense information.
This example suggests that the consequences are more dramatic for a frameshift mutation than for a base-pair substitution. Please keep in mind that the replacement of a cat by a car may also have huge consequences in daily life!

Mutagenic compounds

Base-pair substitutions are often caused by electrophilic substances that want to take up an electron from especially the nucleophilic guanine base that wants to donate an electron to form an electron pair. The consequent guanine addition product (adduct) forms a base-pair with thymine causing a base-pair substitution from G-C to A-T. Alternatively, the guanine adduct may split from the phosphate-sugar backbone of the DNA, leaving an "empty" nucleotide spot in the triplet that can be taken by any nucleotide during DNA replication. Alternatively, base-pair substitutions may be caused by reactive oxygen species (ROS), which are radical compounds that also take up an electron from guanine and form guanine oxidation products (for instance hydroxyl adducts). It should be realized that a DNA adduct can only cause an error in the order of nucleotides (i.e. a mutation) if it is present during DNA replication. Before a cell goes into the DNA synthesis phase of the cell cycle, however, the DNA is thoroughly checked, and possible errors are repaired by DNA repair systems.

Exposure to **direct mutagenic** electrophilic agents rarely occurs because these substances are so reactive that they immediately bind to proteins and DNA in our food and environment. Therefore, DNA damage by such substances in most cases originates from **indirect mutagenic** compounds, which are activated into DNA-binding agents during Phase I of the biotransformation. This process of **bioactivation** is a side-effect of the biotransformation, which is actually aiming at rapid detoxification and elimination of toxic compounds.

Frame-shift mutations are often caused by **intercalating** agents. Unlike electrophilic agents and ROS, intercalating agents do not form covalent bonds with the DNA bases. Instead, due to their planar structure intercalating agents fit exactly between two adjacent nucleotides in the DNA helix. As a consequence, they hinder DNA replication, causing the insertion of an extra nucleotide or the deletion of an original nucleotide in the replicate DNA strain.

Ames test for mutagenicity
Mutagenicity of a compound can be tested in the Ames test, named after Bruce Ames who developed the assay in the early 1970s. The assay makes use of a Salmonella bacteria strain that contains a mutation in a gene encoding for an enzyme involved in the synthesis of the amino acid histidine. Consequently, the bacteria can no longer produce histidine (become "his-") and become auxotrophic, i.e. they depend on their culture medium for histidine. In the assay, the bacteria are exposed to the test compound in a medium that does not contain histidine. If the test compound is not mutagenic, the bacteria cannot grow and will die. If the test compound is mutagenic, it may cause a back-mutation (reversion) of the original mutation in a few bacteria, restoring the autotrophic capacity of the bacteria (i.e. their capacity to produce their own histidine). Growth of mutated bacteria on the histidine depleted medium can be followed by counting colonies (on an agar plate) or by measuring metabolic activity (in a fluctuation assay). Direct mutagenic compounds can be tested in the Ames test without extra treatment. Indirect mutagenic compounds, however, have to be bio-activated before they exert their mutagenic action. For this purpose, a liver homogenate is added to the culture medium containing all enzymes and cofactors required for Phase-I biotransformation of the test compound. This liver homogenate with induced cytochrome P450 (cyp) activity is usually obtained from rats exposed to mixed-type of inducers (i.e. cyp1a, cyp2b, cyp3a), such as the PCB-mixture Aroclor 1254.

Compounds involved in tumour promotion and tumour progression

As stated above, non-mutagenic carcinogens are involved in stimulating the tumour promotion. Tumour promoting substances stimulate cell proliferation and inhibit cell differentiation and apoptosis. Unlike mutagenic compounds, tumour promoting compounds do not interfere directly with DNA and their effect is reversible. Many endogenous substances (e.g. hormones) may act as tumour promoting agents.

The first illustration that chemicals may induce cancer comes from the case of the chimney sweepers in London around 1775. The surgeon Percival Pott (1714-1788) noticed that many adolescent male patients who had developed scrotal cancer had worked during their childhood as a chimney sweeper. Pott made a direct link between exposure to soot during childhood and development of cancer at later age. Based on this discovery, taking a shower after work became mandatory for children working as chimney sweepers, and the observed scrotum cancer incidence decreased. As such, Percival Pott was the first person (i) to link cancer development to chemical substances, (ii) to link early exposure to later cancer development, and (iii) to obtain better occupational health by decreased exposure through better hygiene. In retrospective, we now know that the mutagens involved were polycyclic aromatic hydrocarbons (PAHs) that were bio-activated into highly reactive diol-epoxide metabolites. The delay in cancer development after the early childhood exposure can be attributed to the absence of a tumour promotor. Only after the chimney sweepers had gone through puberty they had sufficient testosterone levels, which stimulates scrotum tissue growth and in this case acted as an endogenous tumour promoting agent.

Tumour progression is the result of aberrant transcriptional activity from either genetic or epigenetic alterations. Genetic alterations can be caused by substances that damage the DNA (called genotoxic substances) and thereby introduce strand breaks and incorrect chromosomal division after mitosis. This results in the typical instable chromosomal characteristics of a malign tumour cell, i.e. a karyotype consisting of reduced and increased numbers of chromosomes (called **aneuploidy** and **polyploidy**, respectively) and damaged chromosomal structures (**abberations**). Chemical substances causing aneuploidy are called aneugens and substances causing chromosomal abberations are called clastogens. Genotoxic substances are also very often mutagenic compounds. Multiple mutations in so-called proto-oncogenes and tumour suppressor genes are necessary to transform a normal cell into a tumour cell. In a healthy cell, cell proliferation is under control by proto-oncogenes that stimulate cell proliferation and tumour suppressor genes that inhibit cell proliferation. In a cancer cell, the balance between proto-oncogenes and tumour suppressor genes is disturbed: proto-oncogenes act as oncogenes, meaning that they continuously stimulate cell proliferation, due to mutations and polyploidy, whereas tumour suppressor genes have become inactive due to mutations and aneuploidy.

Epigenetic alterations are changes in the DNA, but not in its order of nucleotides. Typical epigenetic changes include changes in DNA methylation, histone modifications, and microRNA expression. Compounds that change the epigenome may stimulate tumour progression for instance by stimulating expression of oncogenes and inhibiting expression of tumour suppressor genes. The role in tumour promotion and progression of substances that are capable to induce epigenetic changes is a field of ongoing study.

4.2.7. Question 1

What are the three characteristics of the different cancer development stages?

4.2.7. Question 2

What is the difference between a direct and an indirect mutagenic substance?

4.2.7. Question 3

Explain the principle of the Ames test?

4.2.7. Question 4

What is the difference between a base-pair substitution and a frameshift mutation?

4.2.8. Endocrine disruption

Author: Majorie van Duursen

Reviewer: Timo Hamers, Andreas Kortenkamp

Learning objectives

You should be able to

- explain how xenobiotics can interact with the endocrine system and hormonal actions;
- describe the thyroid system and molecular targets for thyroid hormone disruption;
- explain the concept "it's the timing of the dose that makes the poison".

Keywords: Endocrine system; Endocrine Disrupting Chemical (EDC); DES; Thyroid hormone disruption; Multi- and transgenerational effects

Short history

The endocrine system plays an essential role in the short- and long-term regulation of a variety of biochemical and physiological processes, such as behavior, reproduction, growth as well as nutritional aspects, gut, cardiovascular and kidney function and the response to stress. As a consequence, chemicals that cause changes in hormone secretion or in hormone receptor activity may target many different organs and functions and may result in disorders of the endocrine system and adverse health effects. The nature and the size of endocrine effects caused by chemicals depend on the type of chemical, the level and duration of exposure as well as on the timing of exposure.

The "DES drug disaster" is one of the most striking examples that endocrine-active chemicals can have severe adverse health effects in humans. There was a time when the synthetic estrogen diethylstilbestrol (DES) was considered a miracle drug (Figure 1). DES was prescribed from the 1940s-1970s to millions of women around the world to prevent miscarriages, abortion and premature labor. However, in the early 1970s it was found that daughters of mothers who took DES during their pregnancy have an increased risk of developing a specific vaginal and cervical cancer type. Other studies later demonstrated that women who had been exposed to DES in the womb (in utero) also had other health problems, like increased risk of breast cancer, increased incidence of genital malformations, infertility, miscarriages, and complicated pregnancies. Now, even two generations later, babies are born with reproductive tract malformations that are suspected to be caused by this drug their great grandmothers took during pregnancy. The effects of DES are attributed to the fact that it is a synthetic estrogen (i.e. a xenobiotic compound having similar properties as the natural estrogen 17β-estradiol), thereby disrupting normal endocrine regulation as well as epigenetic processes during development (link to section on Developmental Toxicity).

Around the same time of the DES drug disaster, Rachel Carson wrote a New York Times best-seller called Silent Spring. The book focused on endocrine disruptive properties of persistent environmental contaminants, such as the insecticide DDT (Dichloro Diphenyl Trichloroethane). She wrote that these environmental contaminants were badly degradable in the environment and cause reproductive failure and population decline in a variety of wild life. At the time the book was published, endocrine disruption was a controversial scientific theory that was met with much scepticism as empirical evidence was largely lacking. Still, the book of Rachel Carson has encouraged scientific, societal and political discussions about endocrine disruption. In 1996, another popular

scientific book was published that presented more scientific evidence to warn against the effects of endocrine disruption: Our Stolen Future: Are We Threatening Our Fertility, Intelligence, and Survival? A Scientific Detective Story by Theo Colborn, Dianne Dumanoski and John Peterson Myers.



Figure 1: Advertisement from the 1950s for desPLEX, a synthetic drug containing diethylstilbestrol.

Currently, endocrine disruption is a widely accepted concept and many scientific studies have demonstrated a wide variety of adverse health effects that are attributed to exposure to endocrine active compounds in our environment. Human epidemiological studies have shown dramatic increases in incidences of hormone-related diseases, such as breast, ovarian, testicular and prostate cancer, endometrial diseases, infertility, decreased sperm quality, and metabolic diseases. Considering that hormones play a prominent role in the onset of these diseases, it is highly likely that exposure to endocrine disrupting compounds contributes to these increased disease incidences in humans. In wildlife, the effects of endocrine disruption include feminizing and demasculinizing effects leading to deviant sexual behaviour and reproductive failure in many species, such as fish, frogs, birds and panthers. A striking example of endocrine disruption can be found in the lake Apopka alligator population. Lake Apopka is the third largest lake in the state of Florida, located a few kilometres north west of Orlando. In July 1980, heavy rainfall caused the spill of huge amounts of DDT in the lake by a local pesticide manufacturer. After that, the alligator population in Lake Apopka started to show a dramatic decline. Upon closer examination, these alligators had higher estradiol and lower testosterone levels in their blood, causing poorly developed testes and extremely small penises in the male offspring and severely malformed ovaries in females.

What's in a name: EDC definition

Since the early discussions on endocrine disruption, the World Health Organisation (WHO) has published several reports to present the state-of-the-art in scientific evidence on endocrine disruption, associated adverse health effects and the underlying mechanisms. In 2002, the WHO proposed a definition for an endocrine disrupting compound (EDC), which is still being used. According to the WHO, an EDC can be defined as "an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub) populations." In 2012, WHO stated that "EDCs have the capacity to interfere with tissue and organ development and function, and therefore they may alter susceptibility to different types of diseases throughout life. This is a global threat that needs to be resolved." The European Environment Agency concluded in 2012 that "chemically induced endocrine disruption likely affects human and wildlife endocrine health the world over." A recent report (Demeneix & Slama, 2019) that was commissioned by the European Parliament concluded that the lack of EDC consideration in regulatory procedures is "clearly detrimental for the environment, human health, society, sustainability and most probably for our economy".

The endocrine system

Higher animals, including humans, have developed an endocrine system that allows them to regulate their internal environment. The endocrine system is interconnected and communicates bidirectionally with the neuro- and immunesystems. The endocrine system consist of glands that secrete hormones, the hormones themselves and targets that respond to the hormone. Glands that secrete hormones include the pituitary, thyroid, adrenals, gonads and pancreas. There are three major classes of hormones: amino-acid derived hormones (e.g. thyroid hormones T3 and T4), peptide hormones (e.g. pancreatic hormones) and steroid hormones (e.g. testosterone and estradiol). Hormones elicit a wide variety of biological responses, which almost always start with binding of a hormone to a receptor in its target tissue. This will trigger a chain of intracellular events and eventually a physiological response. Understanding the chemical characteristics of a hormone and its function, may help explain the mechanisms by which chemicals can interact with the endocrine system and subsequently cause adverse health effects.

Mechanism of action

Inherent to the complex nature of the endocrine system, endocrine disruption comes in many shapes and forms. It can occur at the receptor level (link to section on Receptor Interaction), but endocrine disruptors can also disturb the synthesis, metabolism or transport of hormones (locally or throughout the body), or display a combination of multiple mechanisms. For example, DDT can decrease testosterone levels via increased testosterone conversion by the aromatase enzyme, but also acts like an anti-androgen by blocking the androgen receptor and as an estrogen by activating the estrogen receptor. PCBs, polychlorinated biphenyls, are well-characterized thyroid hormone disrupting chemicals. PCBs are industrial chemicals that were widely used in transformators until their ban in the 1970s, but, due to their persistency, PCBs can still ubiquitously be found in the environment, human and wildlife blood and tissue samples (link to section on POPs). PCBs are known to interfere with the thyroid system via inhibition of thyroid hormone synthesis and/or increasing thyroid hormone metabolism, inhibiting binding of thyroid hormones to serum binding proteins, or blocking the ability of thyroid hormones to thyroid hormone receptors. These thyroid disrupting effects can occur in different organs throughout the body (see Figure 1 in Gilbert et al., 2012).

The dose concept

In the 18th Century, physician and alchemist Paracelsus phrased the toxicological paradigm: "Everything is a poison. Only the dose makes that something is not a poison" (link to section Concentration-response relationships, and to Introduction). Generally, this is understood as "the effect of the poison increases with the dose". According to this paradigm, upon deteriming the exposure levels where the toxic response begins and ends, safety levels can be derived to protect humans, animals and their environment. However, interpretation and practical implementation of this concept is challenged by issues that have arisen in modern-day toxicology, especially with EDCs, such as non-monotonic dose-response curves and timing of exposure.

To establish a dose-response relationship, traditionally, toxicological experiments are conducted where adult animals are exposed to very high doses of a chemical. To determine a safe level, you determine the highest test dose at which no toxic effect is seen (the NOAEL or no observed adverse effect level) and add an additional "safety" or "uncertainty" factor of usually 100. This factor 100 accounts for differences between experimental animals and humans, and differences within the human population (see chapter 6 on Risk assessment). Exposures below the safety level are generally considered safe. However, over the past years, studies measuring the effects of hormonally active chemicals also began to show biological effects of endocrine active chemicals at extremely low concentrations, which were presumed to be safe and are in the range of human exposure levels. There are several physiological explanations to this phenomenon. It is important to realize that endogenous hormone responses do not act in a linear, mono-tonic fashion (i.e. the effect goes in one direction), as can be seen in Figure 2 for thyroid hormone levels and IQ. There are feedback loops to regulate the endocrine system in case of over- or understimulation of a receptor and there are clear tissue-differences in receptor expression and sensitivity to hormonal actions. Moreover, hormones are messengers, which are designed to transfer a message across the body. They do this at extremely low concentrations and small changes in hormone concentrations can cause large changes in receptor occupancy and receptor activity. At high concentrations, the change in receptor occupancy is only minimal. This means that the effects at high doses do not always predict the effects of EDCs at lower doses and vice versa.

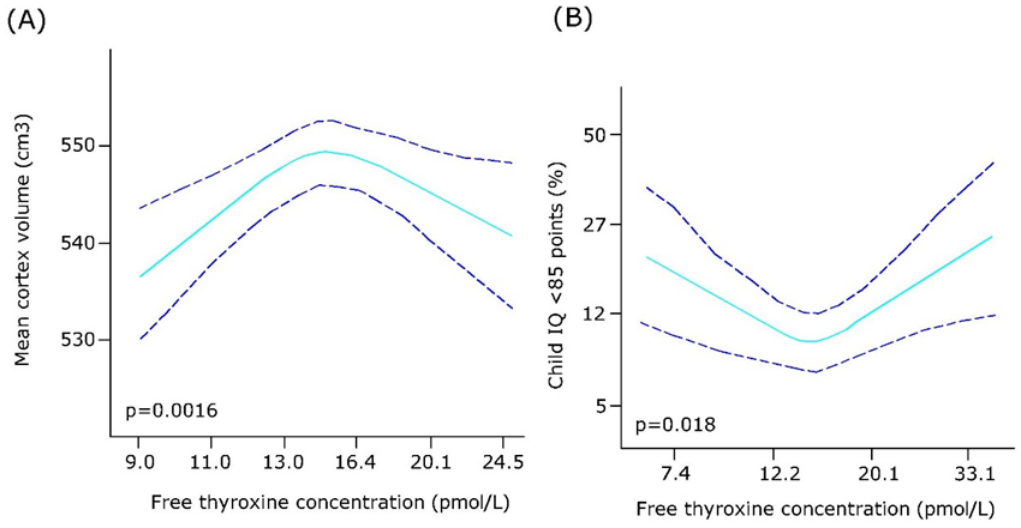


Figure 2. Relation between maternal thyroid hormone level (thyroxine) during pregnancy and (A) offspring cortex volume at the age of 8 years; (B) the predicted probability of offspring having an Intellectual Quotient (IQ) at the age of 6-8 years below 85 points. As women with overt hyperthyroidism or hypothyroidism were excluded, the range of values corresponds to those that can be considered within the normal limits for pregnancy of free thyroxine. Redrawn from Korevaar et al. (2016) by Wilma IJzerman.

It is becoming increasingly clear that not only the dose, but also the timing of exposure plays an important role in determining health effects of EDCs. Multi-generational studies show that EDC exposure in utero can affect future generations (Figure 3). Studies on the grandsons and granddaughters whose mothers were exposed prenatally to DES are limited as they are just beginning to reach the age when relevant health problems, such as fertility, can be studied. However, rodent studies with DES, bisphenol-A and DEHP show that perinatally exposed mothers have grandchildren with malformations of the reproductive tract as well as an increased susceptibility to mammary tumors in female offspring and testicular cancer and poor semen quality in male offspring. Some studies even show effects in the great-grandchildren (F3 generation), which indicates that endocrine disrupting effects have been passed to next generations without direct exposure of these generations. These are called trans-generational effects. Long-term, delayed effects of EDCs are thought to arise from epigenetic modifications in (germ) cells and can be irreversible and transgenerational (link to section on Developmental Toxicity). Consequently, safe levels of EDC exposure may vary, dependent on the *timing* of exposure. Adult exposure to EDCs is often considered activational, e.g. an estrogen-like compounds such as DES can stimulate proliferation of estrogen-sensitive breast cells in an adult leading to breast cancer. When exposure to EDCs occurs during development, the effects are considered to be organizational, e.g. DES changes germ cell development of perinatally exposed mothers and subsequently leads to genital tract malformations in their grandchildren. Multi-generational effects are clear in rodent studies, but are not so clear in humans. This is because it is difficult to characterize EDC exposure in previous generations (which may span over 100 years in humans), and it is challenging to filter out the effect of one specific EDC as humans are exposed to a myriad of chemicals throughout their lives.

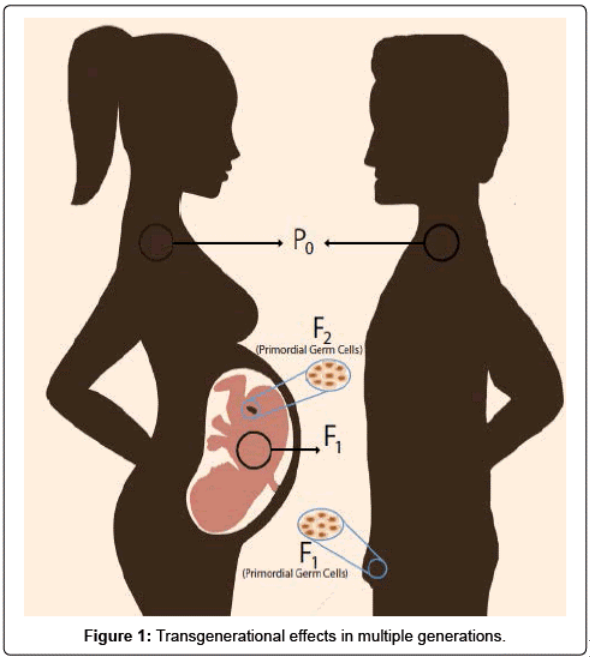


Figure 1: Transgenerational effects in multiple generations.

Figure 3: Exposure to EDCs can affect multiple generations. EDC exposure of parents (P0) can be multi-generational and lead to adverse health effects in children (F1) and grandchildren (F2). Some studies show adverse health effects in great-grandchildren (F3) upon exposure of the parent (P0). This is considered trans-generational, which means that no direct exposure of F3 has taken place, but that effects are passed on via epigenetic modifications in the germ cells of P0, F1 and/or F2. Source: https://www.omicsonline.org/open-access/epigenetic-effects-of-endocrine-disrupting-chemicals-2161-0525-1000381.php?aid=76673

EDCs in the environment

Some well-known examples of EDCs are pesticides (e.g. DDT), plastic softeners (e.g. phthalates, like DEHP), plastic precursors (e.g. bisphenol-A), industrial chemicals (e.g. PCBs), water- and stain-repellents (perfluorinated substances such as PFOS and PFOA) and synthetic hormones (e.g. DES). Exposure to EDCs can occur via air, housedust, leaching into food and feed, waste- and drinking water. Exposure is often unintentional and at low concentrations, except for hormonal drugs. Clearly, synthetic hormones can also have beneficial effects. Hormonal cancers like breast and prostate cancers can be treated with synthetic hormones. And think about the contraceptive pill that has changed the lives of many women around the world since the 1960s. Nowadays, no other method is so widely employed in so many countries around the world as the birth control pill, with an estimate of 75 million users among reproductive-age women with a partner. An unfortunate side effect of this is the increase in hormonal drug levels in our environment leading to feminization of male fish swimming in the polluted waters. Pharmaceutical hormones, along with naturally-produced hormones, are excreted by women and men and these are not fully removed through conventional wastewater treatments. In addition, several pharmaceuticals that are not considered to act via the endocrine system, can in fact display endocrine activity and cause reproductive failure in fish. These are for example the beta-blocker atenolol, antidiabetic drug metformin and analgesic paracetamol.

Further reading:

WHO-IPCS. Global Assessment of the State-of-the-Science of Endocrine Disruptors. 2002 https://www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en/

WHO-UNEP. State of the Science of Endocrine Disrupting Chemicals.2012 www.who.int/iris/bitstream/10665/78101/1/9789241505031_eng.pdf

European Environment Agency (EEA) The impacts of endocrine disrupters on wildlife, people and their environments The Weybridge+15 (1996-2011) report. EEA Technical report No 2/2012 EEA Copenhagen. ISSN 1725-2237 https://www.eea.europa.eu/publications/the-impacts-of-endocrine-disrupters

Demeneix, B., Slama, R. (2019) Endocrine Disruptors: from Scientific Evidence to Human Health Protection. Policy Department for Citizens' Rights and Constitutional Affairs Directorate General for Internal Policies of the Union PE 608.866

http://www.europarl.europa.eu/RegData/etudes/STUD/2019/608866/IPOL_STU(2019)608866_EN.pdf

4.2.8. Question 1

What are possible mechanisms to reduce the action of a certain hormone?

4.2.8. Question 2

Give three target sites for thyroid hormone discuption and name the biological process at that target site that can be affected by an EDC.

4.2.8. Question 3

What mechanism can cause demasculinization of male aligators by DDT?

4.2.8. Question 4

Why is timing of exposure important when assessing the risk of EDC exposure?

4.2.9. Developmental toxicity

Author: Jessica Legradi, Marijke de Cock

Reviewer: Paul Fowler

Learning objectives

You should be able to

- explain the six principles of teratology
- name the difference between deformation, malformation and syndrome
- describe the principle of DoHAD
- indicate what epigenetics is and what epigenetic mechanisms could lead to transgenerational effects

Keywords: Teratogenicity, developmental toxicity, DoHAD, Epigenetics, transgenerational

Developmental toxicity

Developmental toxicity refers to any adverse effects, caused by environmental factors, that interfere with homeostasis, normal growth, differentiation, or development before conception (either parents), during prenatal development, or postnatally until puberty. The effects can be reversible or irreversible. Environmental factors that can have an impact on development are lifestyle factors like alcohol, diet, smoking, drugs, environmental contaminants, or physical factors. Anything that can disturb the development of the embryo or foetus and produces a malformation is called a teratogen. Teratogens can terminate a pregnancy or produce adverse effects called congenital malformations (birth defects, anomaly). A malformation refers to any effect on the structural development of a foetus (e.g. delay, misdirection or arrest of developmental processes). Malformations occur mostly early in development and are permanent. Malformations should not be confused with deformations, which are mostly temporary effects caused by mechanical forces (e.g. moulding of the head after birth). One teratogen can induce several different malformations. All malformations caused by one teratogen are called a syndrome (e.g. fetal alcohol syndrome).

Six Principles of Teratology (by James G. Wilson)

In 1959 James G. Wilson published the 6 principles of teratology. Till now these principles are still seen as the basics of developmental toxicology. The principles are:

1. Susceptibility to teratogenesis depends on the genotype of the conceptus and a manner in which this interacts with adverse environmental factors

Species differences: different species can react different (sensitivities) to the same teratogen. For example, thalidomide (softenon) a drug used to treat morning sickness of pregnant woman causes severe limb malformations in humans whereas such effects were not seen in rats and mice.

Strain and intra litter differences: the genetic background of individuals within one species can cause differences in the response to a teratogen.

Interaction of genome and environment: organisms of the same genetic background can react differently to a teratogen in different environments.

Multifactorial causation: the summary of the above. The severity of a malformation depends on the interplay of several genes (inter and intra species) and several environmental factors.

2. Susceptibility to teratogenesis varies with the developmental stage at the time of exposure to an adverse influence

During development there are periods where the foetus is specifically sensitive to a certain malformation (Figure 1). In general, the very early (embryonic) development is more susceptible to teratogenic effects.

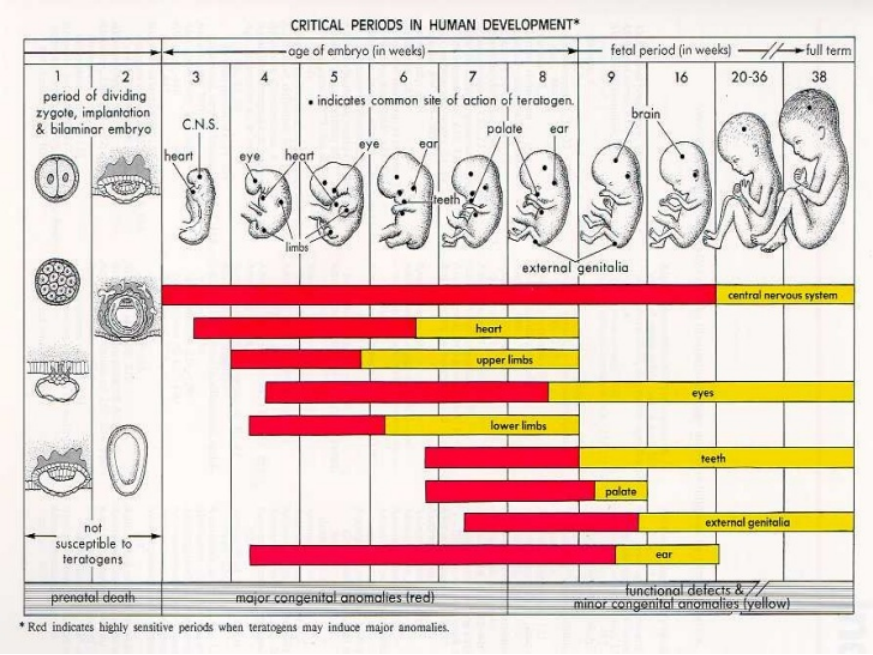


Figure 1: The critical (sensitive) periods during human development. During these periods tissues are more sensitive to malformations when exposed to a teratogen. The timing of period is different for different tissues. Source: https://www.slideshare.net/SDRTL/critical-periods-in-human-development

3. Teratogenic agents act in specific ways (mechanisms) on developing cells and tissues to initiate sequences of abnormal developmental events (pathogenesis)

Every teratogenic agent produces a distinctive malformation pattern. One example is the *foetal alcohol syndrome*, which induces abnormal appearance, short height, low body weight, small head size, poor coordination, low intelligence, behaviour problems, and problems with hearing or seeing and very characteristic facial features (increased distance between the eyes).

4. The access of adverse influences to developing tissues depends on the nature of the influence (agent)

Teratogens can be radiation, infections or chemicals including drugs. The teratogenic effect depends on the concentration of a teratogen that reaches the embryo. The concentration at the embryo is influenced by the maternal absorption, metabolisation and elimination and the time the agent needs to get to the embryo. This can be very different between teratogens. For example, strong radiation is also a strong teratogen as it easily reaches all tissues of the embryo. This also means that a compound tested to be teratogenic in an *in vitro* test with embryos in a tube might not be teratogenic to an embryo in the uterus of a human or mouse as the teratogen may not reach the embryo at a critical concentration.

5. The four manifestations of deviant development are death, malformation, growth retardation and functional deficit

A teratogen can cause minor effects like functional deficits (e.g. reduced IQ), growth retardations or adverse effects like malformations or death. Depending on the timing of exposure and degree of genetic sensitivity, an embryo will have a greater or lesser risk of death or malformations. Very early in development, during the first cell divisions, an embryo will be more likely to die rather than being implanted and developing further.

6. Manifestations of deviant development increase in frequency and degree as dosage increases, from the no-effect to the 100% lethal level

The number of effects and the severity of the effects increases with the concentration of a teratogen. This means that there is a threshold concentration below which no teratogenic effects occur (no effect concentration).

Developmental Origins of Health and Disease (DOHaD)

The concept of Developmental Origins of Health and Disease (DOHaD) describes that environmental factors early in life contribute to health and disease later in life. The basis of this concept was the Barker hypothesis, which was formulated as an explanation for the rise in cardiovascular disease (CVD) related mortality in the United Kingdom between 1900 and 1950. Barker and colleagues observed that the prevalence of CVD and stroke was correlated with neo- and post-natal mortality (Figure 2). This led them to formulate the hypothesis that poor nutrition early in life leads to increased risk of cardiovascular disease and stroke later in life. Later, this was developed into the thrifty phenotype hypothesis stating that poor nutrition in utero programs for adaptive mechanisms that allow to deal with nutrient-poor conditions in later life, but may also result in greater susceptibility to metabolic syndrome. This thrifty hypothesis was finally developed into the DOHaD theory.

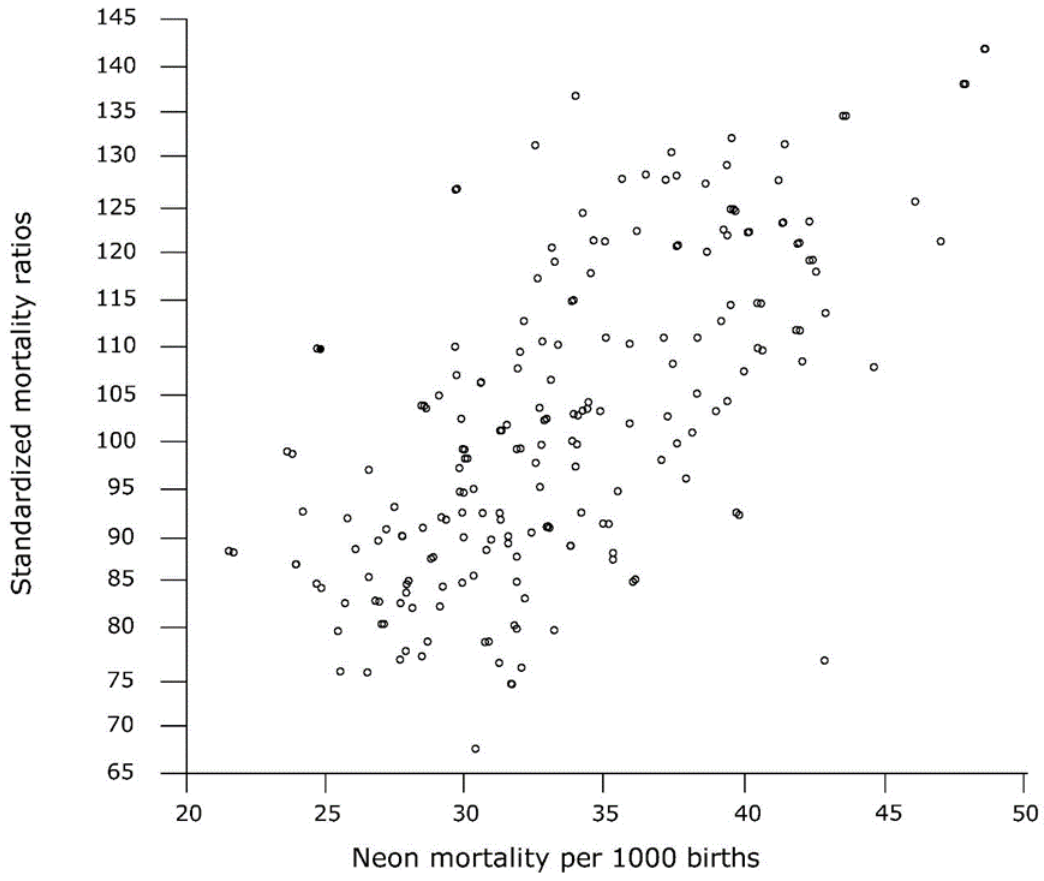


Figure 2:
Standardized mortality ratios for ischaemic heart disease in both sexes (y-axis) and neonatal mortality per 1000 births, 1921-1925 (x-axis). Redrawn from Barker et al. (1986) by Wilma Ijzerman.

The effect of early life nutrition on adult health is clearly illustrated by the Dutch Famine Birth Cohort Study. In this cohort, women and men who were born during or just after the Dutch famine, were studied retrospectively. The Dutch famine was a famine that took place in the Western part of the German-occupied Netherlands in the winter of 1944-1945. Its 3-months duration creates the possibility to study the effect of poor nutrition during each individual trimester of pregnancy. Effects on birth weight, for example, may be expected if caloric intake during pregnancy is restricted. This was, however, only the case when the famine occurred during the second or the third trimesters. Higher glucose and insulin levels in adulthood were only seen for those exposed in the third trimester, whereas those exposed during the second trimester showed a higher prevalence of obstructive airways disease. These effects were not observed for the other trimesters, which can be explained by the timing of caloric restriction during pregnancy: during normal pregnancy pancreatic islets develop during the third trimester, while during the second trimester the number of lung cells is known to double.

The DOHaD concept does not merely focus on early life nutrition, but includes all kinds of environmental stressors during the developmental period that may contribute to adult disease, including exposure to chemical compounds. Chemicals may elicit effects such as endocrine disruption or neurotoxicity, which can lead to permanent morphological and physiological changes when occurring early in life. Well-known examples of such chemicals are diethylstilbestrol (DES) and dichlorodiphenyltrichloroethane (DDT). DES was an estrogenic drug given to women between 1940 and 1970 to prevent miscarriage. It was withdrawn from the market in 1971 because of carcinogenic effects as well as an increased risk for infertility in children who were exposed in utero (link to section on Endocrine disruption). DDT is a pesticide that has been banned in most countries, but is still used in some for malaria control. Several studies, including a pooled analysis of seven European cohorts, found associations between in utero DDT exposure levels and infant growth and obesity.

The ubiquitous presence of chemicals in the environment makes it extremely relevant to study health effects in humans, but also makes it very challenging as virtually no perfect control group exists. This emphasizes the importance of prevention, which is the key message of the DOHaD concept. Adult lifestyle and corresponding exposure to toxic compounds remain important modifiable factors for both treatment and prevention of disease. However, as developmental plasticity, and therefore the potential for change, is highest early in life, it is important to focus on exposure in the early phases: during pregnancy, infancy, childhood and adolescence. This is reflected by regulators frequently imposing lower tolerable exposure levels for infants compared to adults.

Epigenetics

For some compounds in utero exposure is known to cause effects later in life (see DOHad) or even induce effects in the offspring or grand-offspring of the exposed embryo (transgenerational effect). Diethylstilbesterol (DES) is a compound for which transgenerational effects are reported. DES was given to pregnant women to reduce the risk of spontaneous abortions and other pregnancy complications. Women who took DES during pregnancy have a slightly increased risk of breast cancer. Daughters exposed in utero, on the other hand, had a high tendency to develop rare vaginal tumours. In the third-generation, higher incidences of infertility, ovarian cancer and an increased risk of birth defects were observed. However, the data available for the third generation is small and therefore possess only limited evidence so far. Another compound suspected to cause transgenerational effects is the fungicide vinclozolin. Vinclozolin is an anti-androgenic endocrine disrupting chemical. It has been shown that exposure to vinclozolin leads to transgenerational effects on testis function in mice.

Transgenerational effects can be induced via genetic alterations (mutations) in the DNA. Thereby the order of nucleotides in the genome of the parental gametocyte is altered and this alteration is inherited to the offspring. Alternatively, transgenerational effects can be induced is via epigenetic alterations. Epigenetics is defined as the study of changes in gene expression that occur without changes in the DNA sequence, and which are heritable in the progeny of cells or organisms. Epigenetic changes occur naturally but can also be influenced by lifestyle factors or diseases or environmental contaminants. Epigenetic alterations are a special form of developmental toxicology as effects might not cause immediate teratogenic malformations. Instead the effects may be visible only later in life or in subsequent generations. It is assumed that compounds can induce epigenetic changes and thereby cause transgenerational effects. For DES and vinclozolin epigenetic changes in mice have been reported and these might explain the transgenerational changes seen in humans. Two main epigenetic mechanisms are generally described as being responsible for transgenerational effects, i.e. DNA methylation and histone modifications.

DNA methylation

DNA methylation is the most studied epigenetic modification and describes the methylation of cytosine nucleotides in the genome (Figure 3) by DNA methyltransferase (DNMTs). Gene activity generally depends on the degree of methylation of the promotor region: if the promotor is methylated the gene is usually repressed. One peculiarity of DNA methylation is that it can be wiped and replaced again during epigenetic reprogramming events to set up cell- and tissue-specific gene expression patterns. Epigenetic reprogramming occurs very early in development. During this phase epigenetic marks, like methylation marks are erased and remodelled. Epigenetic reprogramming is necessary as maternal and paternal genomes are differentially marked and must be reprogrammed to assure proper development.

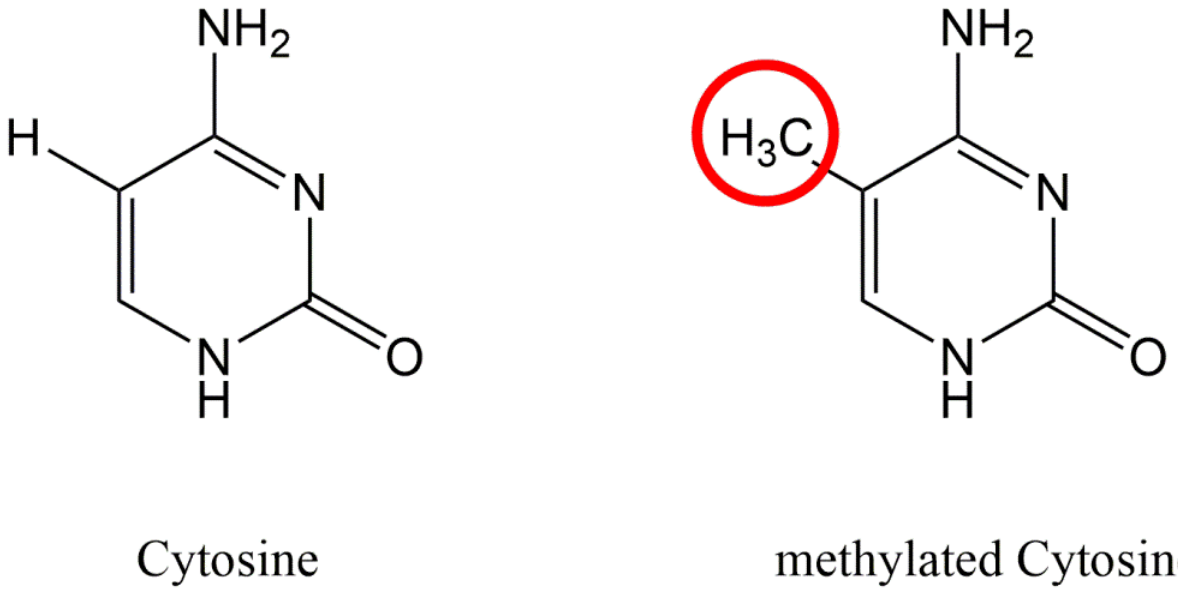


Figure 3: The methylation of the cytosine nucleotide. One hydrogen atom is replaced by a methyl group. Drawn by Steven Droge.

Histone modification

Within the chromosome the DNA is densely packed around histone proteins. Gene transcription can only take place if the DNA packaging around the histones is loosened. Several histone modification processes are involved in loosening this packaging, such as acetylation, methylation, phosphorylation or ubiquitination of the histone molecules (Figure 4).

Figure in preparation

Figure 4: *(a) The DNA is wrapped around the histone molecules. The histone molecules are arranged in a way that their amino acid tails are pointing out of the package. These tails can be altered for example via acetylation. (b) If the tails are acetylated the DNA is packed less tightly and genes can be transcribed. If the tails are not acetylated the DNA is packed very tight and gene transcription is hampered.*

References

Barker, D.J., Osmond, C. (1986). Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. Lancet 1 (8489), 1077-1081.

4.2.9. Question 1

Describe the difference between malformation and deformation?

4.2.9. Question 2

What factors can influence the susceptibility of an organism to teratogen?

4.2.9. Question 3

Describe how DOHaD differs from the Barker hypothesis?

4.2.9. Question 4

What are the two epigenetic mechanisms mostly responsible for transgenerational effects?

4.2.9. Question 5

Name a teratogenic compound?

4.2.10. Immunotoxicity

Author: Nico van den Brink

Review: Manuel E. Ortiz-Santaliestra

Learning objectives:

You should be able to:

- Understand the complexity of potential effects of chemicals on the immune system
- Explain the difference between innate and acquired parts of the immune system
- Explain the most important modes of toxicity that may affect immune cells

Key words: Immune toxicology, pathogens, innate and adaptive immune system, lyme disease

Introduction

The immune system of organisms is very complex with different cells and other components interacting with each other. The immune system has the function to protect the organism from pathogens and infections. It consists of an innate part, which is active from infancy and an acquired part which is adaptive to exposure to pathogens. The immune system may include different components depending on the species (Figure 1).

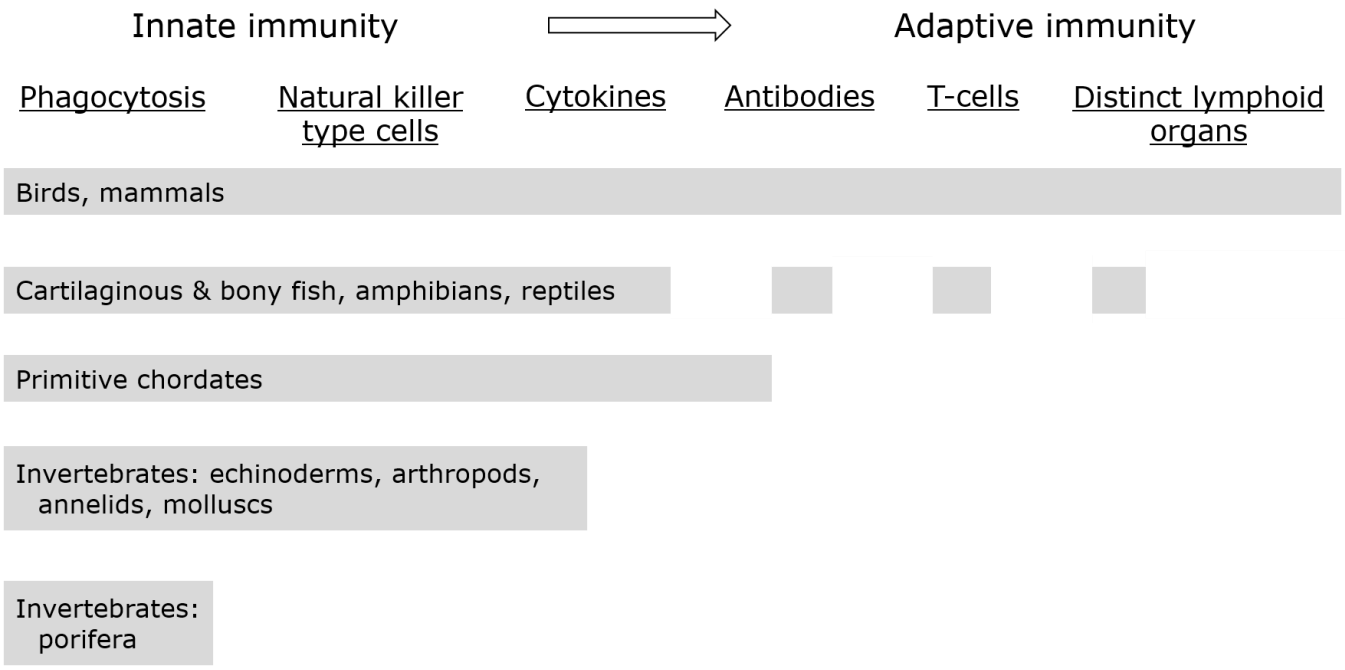


Figure 1. Simplified diagram of the evolution of the immune system indicating some preserved key immunological functions (adapted from Galloway and Handy, 2003).

The main organs involved in the immune system of mammals are spleen, thymus, bone marrow and lymph nodes. In birds, besides all of the above, there is also the bursa of Fabricius. These organs all play specific roles in the immune defence, e.g. the spleen synthesises antibodies and plays an important role in the dynamics of monocytes; the thymus is the organ where T-cells develop while in bone marrow lymphoid cells are produced, which are transported to other tissues for further development. The bursa of Fabricius is specific for birds and is essential for B-cell development. Blood is an important tissue to be considered because of its

role in transporting cells. The innate system generally provides the first response to infections and pathogens, however it is not very specific. It consists of several cell types with different functions like macrophages, neutrophils and mast cells. **Macrophages** and **neutrophils** may act against pathogens by phagocytosis (engulfing in cellular lysosomes and destruction of the pathogen). **Neutrophils** are relatively short lived, act fast and can produce a respiratory burst to destroy the pathogen/microbe. This involves a rapid production of Reactive Oxygen Species (ROS) which may destroy the pathogens. **Macrophages** generally have a longer live span, react slower but more prolonged and may attack via production of nitric oxide and less via ROS. Macrophages produce cytokines to communicate with other members of the immune system, especially cell types of the acquired system. Other members of the innate immune system are mast cells which can excrete e.g. histamine on detection of antigens. Cells of the acquired, or adaptive immune system mount more specific responses for the immune insult, and are therefore generally more effective. **Lymphocytes** are the cells of the adaptive immune system which can be classified in **B-lymphocytes** and **T-lymphocytes**. **B-lymphocytes** produce antibodies which can serve as cell surface antigen-receptors, essential in the recognition of e.g. microbes. **B-lymphocytes** facilitate humoral (extracellular) immune responses against extracellular microbes (in the respiratory gastrointestinal tract and in the blood/lymph circulation). Upon recognition of an antigen, B-lymphocytes produce species antibodies which bind to the specific antigen. This on the one hand may decrease the infectivity of pathogens (e.g. microbes, viruses) directly, but also mark them for recognition by phagocytic cells. **T-lymphocytes** are active against intracellular pathogens and microbes. Once inside cells, pathogens are out of reach of the B-lymphocytes. **T-lymphocytes** may activate macrophages or neutrophils to destroy phagocytosed pathogens or even destroy infected cells. Both B- and T-lymphocytes are capable of producing an extreme diversity of clones, specific for antigen recognition. Communication between the different immune cells occurs by the production of e.g. cytokines, including interleukins (ILs), chemokines, interferons (IFs), and also Tumour Necrosis Factors (TNFs). Cytokines and TNFs are related to specific responses in the immune system, for instance IL6 is involved in activating B-cells to produce immunoglobulins, while TNF-α is involved in the early onset of inflammation, therefore one of the cytokines inducing acute immune responses. Inflammation is a generic response to pathogens mounted by cells of the innate part of the immune system. It generally results in increased temperature and swelling of the affected tissue, caused by the infiltration of the tissue by leukocytes and other cells of the innate system. A proper acute inflammatory response is not only essential as a first defence but will also facilitate the activation of the adaptive immune system. Communication between immune cells, via cytokines, not only directs cells to the place of infection but also activates for instance cells of the acquired immune system. This is a very short and non-exhaustive description of the immune system, for more details on the functioning of the immune system see for instance Abbas et al. (2018).

Chemicals may affect the immune system in different ways. Exposure to lead for instance may result in immune suppression in waterfowl and raptors (Fairbrother et al. 2004, Vallverdú-Coll et al., 2019). Decreasing spleen weights, lower numbers of white blood cells and reduced ability to mount a humoral response against a specific antigen (e.g. sheep red blood cells), indicated a lower potential of exposed birds to mount proper immune responses upon infection. Exposure to mercury resulted in decreased proliferation of B-cells in zebra finches (*Taeniopygia guttata*), affecting the acquired part of the immune system (Lewis et al., 2013). However, augmentation of the immune system upon exposure to e.g. cadmium has also been reported in for instance small mammals, indicating an enhancement of the immune response (Demenesku et al., 2014). Both immune suppression as well as immune enhancement may have negative impacts on the organisms involved; the former may decrease the ability of the organism to deal with pathogens or other infections, while immune enhancement may increase the energy demands of the organism and it may also result in for instance hypersensitivity or even auto-immunity in organisms.

Chemicals may affect immune cells via **toxicity** **to mechanisms that are not specific to the immune system**. Since many different cell types are involved in the immune system, the sensitivity to these modes of toxicity may vary considerably among cells and among chemicals. This would imply that as a whole, the immune system may inherently include cells that are sensitive to different chemicals, and as such may be quite sensitive to a range of toxicants. For instance induction of **apoptosis**, programmed cell death, is essential to clear the activated cells involved in an immune response after the infection is minimised and the system is returning to a state of homeostasis (see Figure 2). Chemicals may induce apoptosis, and thus interfere with the kinetics of adaptive immune responses, potentially reducing the longevity of cells.

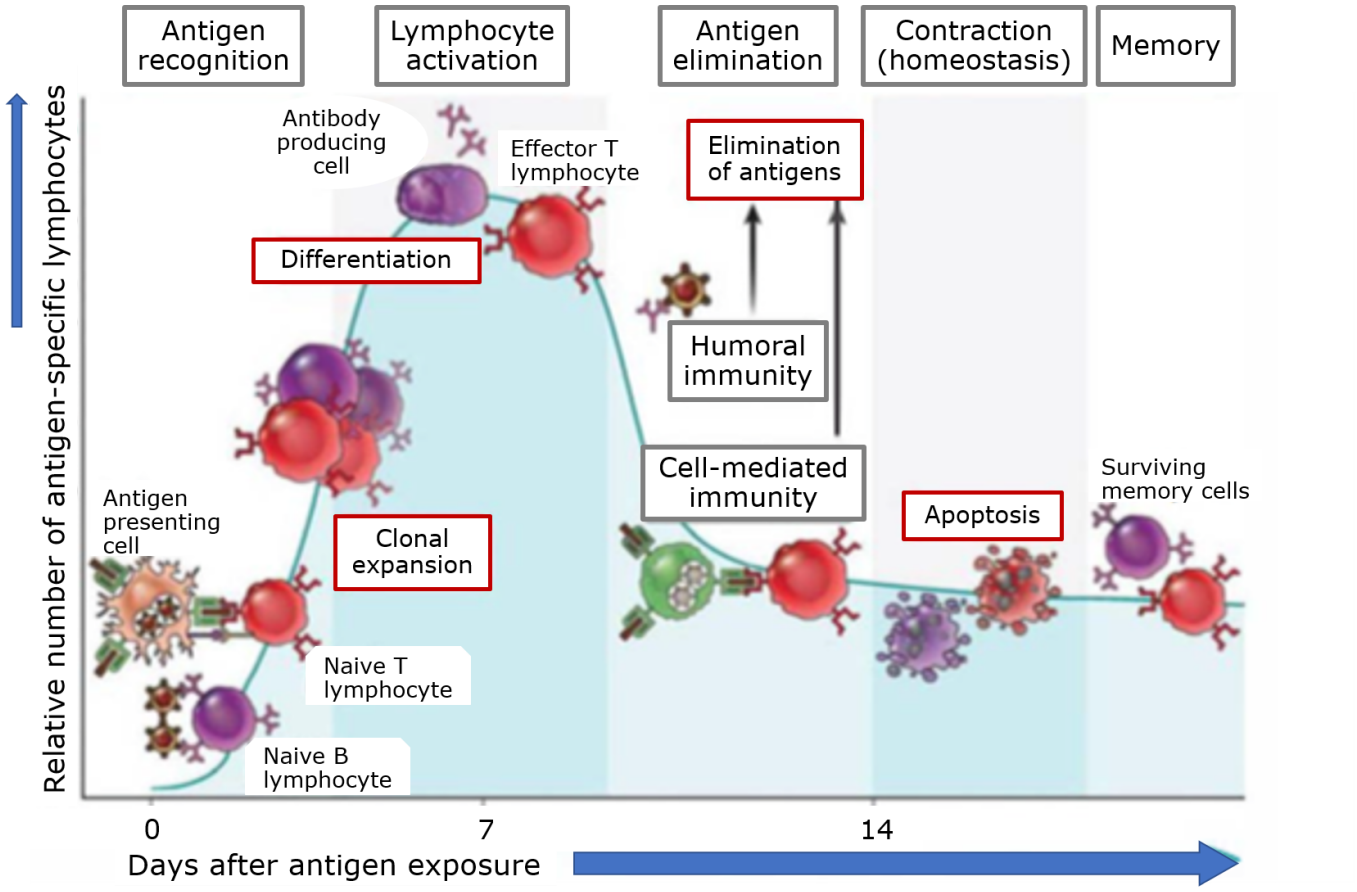


Figure 2. Development of an adaptive immune response, with the different cell types involved. Adapted from www.memorangapp.com/flashcards/170860/Immunology+Exam+1/.

Toxic effects on mechanisms specific to the immune system may be related to its functioning. Since the production of ROS and nitric oxides are effector pathways along which neutrophils and macrophages of the innate systems combat pathogens (via a high production of reactive oxygen species, i.e. oxidative burst, to attack pathogens), impacts on the **oxidative status** of these cells may not only result in general toxicity, potentially affecting a range of cell types, but it may also affect the responsiveness of the (innate) immune system particularly. For instance, cadmium has a high affinity to bind to glutathione (GSH), a prominent anti-oxidant in cells, and has shown to affect acute immune responses in thymus and spleens of mice (Pathak and Khandelwal, 2007) via this mechanism. A decrease of GSH by binding of chemicals (like cadmium) may modulate macrophages towards a pro-inflammatory response by changes in the redox status of the cells involved, changing not only their activities against pathogens but potentially also their production and release of cytokines (Dong et al., 1998).

GSH is also involved in the modulation of the acquired immune system by affecting so-called antigen-presenting cells (APCs, e.g. dendritic cells). APCs capture microbial antigens that enter the body, transport these to specific immune-active tissues (e.g. lymph nodes) and present them to naive T-lymphocytes, inducing a proper immune response, so-called T-helper cells. T-helper cells include subsets, e.g. T-helper 1 cells (Th1-cells) and T-helper 2 cells (Th2-cells). Th1 responses are important in the defence against intracellular infections, by activation of macrophages to ingest microbes. Th2-responses may be initiated by infections by organisms too large to be phagocytosed, and mediated by e.g. allergens. As mentioned, GSH depletion may result in changes in cytokine production by APC (Dong et al., 1998), generally affecting the release of Th1-response promoting cytokines. Exposure to chemicals interfering with GSH kinetics may therefore result in a dis-balance between Th1 and Th2 responses and as such affect the responsiveness of the immune system. Cadmium and other metals have a high affinity to bind to GSH and may therefore reduce Th1 responses, while in contrast, GSH promoting chemicals may reduce the organisms' ability to initiate Th2-responses (Pathak & Khandelwal, 2008).

The overview on potential effects that chemicals may have on the immune system as presented here is not exhaustive at all. This is even more complicated because effects may be contextual, meaning that chemicals may have different impacts depending on the

situation an organism is in. For instance, the magnitude of immunotoxic effects may be dependent on the general condition of the organism, and hence some infected animals may show effects from chemical exposure while others may not. Impacts may also differ between types of infection (e.g. Th1 versus Th2 responsive infections). This, together with the complex and dynamic composition of the immune system, limits the development of general dose response relationships and hazard predictions for chemicals. Furthermore, most of the research on effects of chemicals on the immune system is focussed on humans, based on studies on rats and mice. Little is known on differences among species, especially in non-mammalian species which may have completely differentially structured immune systems. Some studies on wildlife have shown effects of trace metals on small mammals (Tersago et al., 2004, Rogival et al., 2006, Tête et al., 2015) and of lead on birds (Vallverdú-Coll et al., 2015). However, specific modes of action are still to be resolved under field conditions. Research on immuno-toxicity in wildlife however is essential not only from a conservational point of view (to protect the organisms and species involved) but also from the perspective of human health. Wildlife plays an important role in the kinetics of zoonotic diseases, for instance small mammals are the prime reservoir for *Borrelia spirochetes*, the causative pathogens of Lyme-disease while migrating waterfowl are indicated to drive the spread of e.g. avian influenza. The role of wildlife in the kinetics of the environmental spread of zoonotic diseases is therefore eminent, which may seriously be affected by chemical induced alterations of their immune system.

References and further reading

Abbas, A.K., Lichtman, A.H., Pillai, S. (2018). Cellular and Molecular Immunology. 9th Edition. Elsevier, Philadelphia, USA. ISBN: 978-0-323-52324-0

Demenesku, J., Mirkov, I., Ninkov, M., Popov Aleksandrov, A., Zolotarevski, L., Kataranovski, D., Kataranovski, M. (2014). Acute cadmium administration to rats exerts both immunosuppressive and proinflammatory effects in spleen. Toxicology 326, 96-108.

Dong, W., Simeonova, P.P., Gallucci, R., Matheson, J., Flood, L., Wang, S., Hubbs, A., Luster, M.I. (1998). Toxic metals stimulate inflammatory cytokines in hepatocytes through oxidative stress mechanisms. Toxicology and Applied Pharmacology 151, 359-366.

Fairbrother, A., Smits, J., Grasman, K.A. (2004). Avian immunotoxicology. Journal of Toxicology and Environmental Health, Part B 7, 105-137.

Galloway, T., Handy, R. (2003). Immunotoxicity of organophosphorous pesticides. Ecotoxicology 12, 345-363.

Lewis, C.A., Cristol, D.A., Swaddle, J.P., Varian-Ramos, C.W., Zwollo, P. (2013). Decreased immune response in Zebra Finches exposed to sublethal doses of mercury. Archives of Environmental Contamination & Toxicology 64, 327-336.

Pathak, N., Khandelwal, S. (2007). Role of oxidative stress and apoptosis in cadmium induced thymic atrophy and splenomegaly in mice. Toxicology Letters 169, 95-108.

Pathak, N., Khandelwal, S. (2008). Impact of cadmium in T lymphocyte subsets and cytokine expression: differential regulation by oxidative stress and apoptosis. Biometals 21, 179-187.

Rogival, D., Scheirs, J., De Coen, W., Verhagen, R., Blust, R. (2006). Metal blood levels and hematological characteristics in wood mice (*Apodemus sylvaticus* L.) along a metal pollution gradient. Environmental Toxicology & Chemistry 25, 149-157.

Tersago, K., De Coen, W., Scheirs, J., Vermeulen, K., Blust, R., Van Bockstaele, D., Verhagen, R. (2004). Immunotoxicology in wood mice along a heavy metal pollution gradient. Environmental Pollution 132, 385-394.

Tête, N., Afonso, E., Bouguerra, G., Scheifler, R. (2015). Blood parameters as biomarkers of cadmium and lead exposure and effects in wild wood mice (*Apodemus sylvaticus*) living along a pollution gradient. Chemosphere 138, 940-946.

Vallverdú-Coll, N., López-Antia, A., Martinez-Haro, M., Ortiz-Santaliestra, M.E., Mateo, R. (2015). Altered immune response in mallard ducklings exposed to lead through maternal transfer in the wild. Environmental Pollution 205, 350-356.

Vallverdú-Coll, N., Mateo, R., Mougeot, F., Ortiz-Santaliestra, M.E. (2019). Immunotoxic effects of lead on birds. Science of the Total Environment 689, 505-515.

4.2.10. Question 1

Name two general reasons why immunomodulation in organisms may be very sensitive to exposure to environmental chemicals

---

4.2.10. Question 2

Why is it that immunomodulatory chemicals to which humans are not exposed still may have an impact on human health?

4.2.11. Toxicity mechanisms of metals

Author: Nico M. van Straalen

Reviewers: Philip S. Rainbow, Henk Schat

Learning objectives

You should be able to

- list five biochemical categories of metal toxicity mechanisms and describe an example for each case
- interpret biochemical symptoms of metal toxicity (e.g. functional categories of gene expression profiles) and explain these in terms of the mode of action of a particular metal

Keywords: Reactive oxygen species, protein binding, DNA binding, ion pumps,

Synopsis

Toxicity of metals on the biochemical level is due to a wide variety of mechanisms, which may be classified as follows, although they are not mutually exclusive: (1) generation of radical oxygen species (Fe, Cu), (2) binding to nucleophilic groups in proteins (Cd, Pb), (3) binding to DNA (Cr, Cd), (4) binding to ion channels or membrane pumps (Pb, Cd), (5) interaction with the function of essential cellular moieties such as phosphate, sulfhydryl groups, iron or calcium (As, Cd, Al, Pb). In addition, these mechanisms may act simultaneously and interact with each other. There are interesting species patterns of susceptibility to metals, e.g. mammals are hardly susceptible to zinc, while plants and crustaceans are. Earthworms, gastropods and fungi are quite sensitive to copper, but not so for terrestrial vertebrates. In this section we discuss five different categories of metal toxicity as well as some patterns of species differences in sensitivity to metals.

Generation of reactive oxygen species

Reactive oxygen species (ROS) are activated forms of oxygen that have one or more unpaired electrons in the outer orbit. The best knowns are superoxide anion (O_2^-), singlet oxygen ($1\Delta gO_2$), hydrogen peroxide (H_2O_2) and hydroxyl radical ($OH^\bullet$) (see the section on Oxidative stress), effective catalyzers of reactive oxygen species. This relates to their capacity to engage in redox reactions with transfer of one electron. One of the most famous reactions is the so-called Fenton reaction catalyzed by reduced iron and copper ions:

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^\bullet + OH^-$$

$$Cu^+ + H_2O_2 \rightarrow Cu^{2+} + OH^\bullet + OH^-$$

Both reactions produce the highly reactive hydroxyl radical ($OH^\bullet$), which may trigger severe cellular damage by peroxidation of membrane lipids (see the section on Oxidative Stress). Very low concentrations of metal ions can keep this reaction running, because the reduced forms of the metal ions are restored by a second reaction with hydrogen peroxide:

$$Fe^{3+} + H_2O_2 \rightarrow Fe^{2+} + O_2^- + 2H^+$$

$$Cu^{2+} + H_2O_2 \rightarrow Cu^+ + O_2^- + 2H^+$$

The overall reaction is a metal-catalyzed degradation of hydrogen peroxide, causing superoxide anion and hydroxyl radical as intermediates. Oxidative stress is one of the most important mechanisms of toxicity of metals. This can also be deduced from the metal-induced transcriptome. Gene expression profiling has shown that it is not uncommon that more than 10% of the genome responds to sublethal concentrations of cadmium.

Protein binding

Several metals have a great affinity towards sulfhydryl (-SH) groups in the cysteine residues of proteins. Binding to such groups may distort the secondary structure of a protein at sites where SH-groups coordinate to form S-S bridges. The SH-group is a typical example of a nucleophile, that is, a group that easily donates an electron pair to form a chemical bond. The group that accepts the electron pair is called an electrophile. Another amino acid in a protein to which metals are preferentially bound is the imidazole

side-chain of histidine. This heterocyclic aromatic group with two nitrogen atoms easily engages into chemical bonds with metal ions. In fact, histidine residues are often used in metalloproteins to coordinate metals at the active site and to transport metals from the roots upwards through the xylem vessels of plants.

A classical case of metal-protein interaction with subsequent toxicity is the case of lead binding to δ-aminolevulinic acid dehydratase (δ-ALAD). This is an enzyme involved in the synthesis of hemoglobin. It catalyzes the second step in the biosynthetic pathway, the condensation of two molecules of δ-aminolevulinic acid to one molecule of porphobilinogen, which is a precursor of porphyrin, a functional unit binding iron in hemoglobin (Figure 1). The enzyme has several sulfhydryl groups that are susceptible to lead. In the erythrocyte more than 80% of lead is in fact bound to the δ-ALAD protein (much more than is bound to hemoglobin). Inhibition of δ-ALAD leads to decreased porphyrin synthesis, insufficient hemoglobin, loss of oxygen uptake capacity, and eventually anemia.

Because the inhibition of δ-ALAD by lead occurs at already very low exposure levels, it makes a very good biomarker for lead exposure. Measurement of δ-ALAD activity in blood has been conducted extensively in workers of metal-processing industries and people living in metal-polluted environments. Also in fish, birds and several invertebrates (earthworms, planarians) the δ-ALAD assay has been shown to be a useful biomarker of lead exposure. In addition to lead, mercury is known to inhibit δ-ALAD, while the inhibitions by both lead and mercury can be alleviated to some extent by zinc.

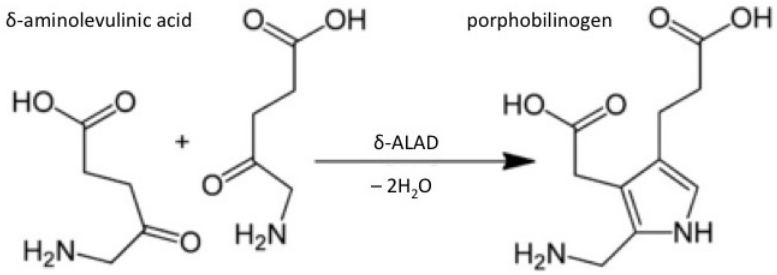


Figure 1. Formation of porphobilinogen from δ-ALA, catalyzed by δ-ALAD.

DNA binding

Chromium, especially the trivalent (Cr^{3+}) and the hexavalent (Cr^{6+}) ions are the most notorious metal species known to bind to DNA. Both trivalent and hexavalent chromium may cause mutations and hexavalent chromium is also a known carcinogen. Although the salts of Cr^{6+} are only slightly soluble, the reactivity of the Cr^{6+}-ion is so pronounced that only very little hexavalent chromium salt is already dangerous.

The genotoxicity of trivalent chromium is due to the formation of crosslinks between proteins and DNA. Any DNA molecule is surrounded by proteins (histones, regulatory proteins, chromatine). Cr^{3+} binds to amino acids such as cysteine, histidine and glutamic acid on the one hand, and to the phosphate groups in DNA on the other, without any preference for a specific nucleotide (base). The result is a covalent bond between DNA and a protein that will inhibit transcription or regulatory functions of the DNA segment involved.

Another metal known to interact with DNA is nickel. Although the primary effects of nickel are to induce allergic reactions, it is also a known carcinogen. The exact molecular mechanism is not as well known as in the case of chromium. Nickel could crosslink proteins and DNA in the same way as chromium, but is also argued that nickel's carcinogenicity is due to oxidative stress, resulting in DNA damage. Another suggested mechanism is that nickel could interfere with the DNA repair system.

Inhibition of ion pumps

Many metals may compete with essential metals during uptake or transport across membranes. A well-known case is the competition between calcium and cadmium at the Ca^{2+}ATPase pump in the basal membrane of fish gills (Figure 2).

The gills of fish serve as a target for many water-born toxic compounds because of their large contact area with the water, consisting of several membranes, each with infoldings to increase the surface area, and also their high metabolic activity which stems from their important regulatory activities (uptake of oxygen, uptake of nutrients and osmoregulation). The single-layered epithelium has two types of cells, one active in osmoregulation (called chloride cells), and one active in transport of nutrients and

oxygen (called respiratory cells). There are strong tight junctions between these cells to ensure complete impermeability of the epithelium to ions. The apical membrane of the respiratory cells has many uptake pumps and channels (Figure 2). Calcium enters the cells though a calcium channel (without energetic costs, following the gradient). The intracellular calcium concentration is regulated by a calcium-ATPase in the basal membrane, which pumps calcium out of the epithelial cells into the blood.

Figure in preparation

Figure 2. *Schematic representation of the cells in a fish gill epithelium, showing the fluxes of calcium and cadmium. Calcium enters the cell through calcium channels on the apical side, and is pumped out of the cells into the circulation by a calcium ATPase in the basal membrane. Cadmium ions enter the cells also through the calcium channels, but inhibit the basal calcium ATPAse, causing hypocalcemia in the rest of the body. m = mucosa (apical side), s = serosa (basal side), BP = binding protein, mito = mitochondrion, ER = endoplasmic reticulum. From Verbost et al. (1989).*

Water-borne cadmium ions, which resemble calcium ions in their atomic radius, enter the cell through the same apical calcium channels, but subsequently inhibit the basal membrane calcium transporter by direct competition with calcium for the binding site on the ATPAse. The consequence is an accumulation of calcium in the respiratory cells, and a lack of calcium in the body of the fish, which causes a variety of secondary effects; amongst others hormonal disturbance, while a severe decline of plasma calcium is a direct cause of mortality. This effect of cadmium occurs at very low concentrations (nanomolar range), and it explains the high toxicity of this metal to fish. Similar cadmium-induced hypocalcemia mechanisms are present in the gill membranes of crustaceans and most likely also in gut epithelium cells of many other species.

Interaction with essential cellular constituents

There are various cellular ligands outside proteins or DNA that may bind metals. Among these are organic acids (malate, citrate), free amino acids (histidine, cysteine), and glutathione. Metals may also interfere with the cellular functions of phosphate, iron, calcium or zinc, for example by replacing these elements from their normal binding sites in enzymes or other molecules. To illustrate a case of interaction with phosphate we discuss shortly the toxicity of arsenic. Arsenic is strictly speaking not a metal, since arsenic oxide may engage in both base-forming and acid-forming reactions. Together with antimony and four other, lesser-known elements, arsenic is indicated as a "metalloid".

Arsenic is a potent toxicant; arsenic trioxide (As_2O_3) is well known for its high mammalian toxicity and its use as a rodenticide and wood preservative. There are also therapeutic applications of arsenic trioxide, against certain leukemias and arsenic is often implied in homeopathic treatments. Arsenic compounds are easily transported throughout the body, also across the placental barrier in pregnant women.

Arsenic can occur in two different valency states: arsenate (As^{5+}) and arsenite (As^{3+}). The terms are also used to indicate the oxy-salts, such as ferric arsenate, $FeAsO_4$, and ferric arsenite, $FeAsO_3$. Inside the body, arsenic may be present in oxidized as well as reduced state, depending on the conditions in the cell, and it is enzymatically converted to one or the other state by reductases and oxidases. It may also be methylated by methyltransferases. The two different forms of arsenic have quite different toxicity mechanisms. Arsenate, AsO_4^{3-}, is a powerful analog of phosphate, while arsenite (AsO_3^{3-}) reacts with SH-groups in proteins, like the metals discussed above. Arsenite is also a known carcinogen; the mechanism seems not to rely on DNA binding, like in the case of chromium, but on the induction of oxidative stress and interference with cellular signaling.

The most common reason of chronic arsenic poisoning is due to inhibition of the enzyme glyceraldehyde phosphate dehydrogenase (GAPDH). This is a critical enzyme of the glycolysis, converting glyceraldehyde-3-phosphate into 1,3-biphosphoglycerate. However, in the presence of arsenate, GAPDH converts glyceraldehyde-3-phosphate into 1-arseno-3-phosphoglycerate. Actually arsenate acts as a phosphate analog to "fool" the enzyme. The product 1-arseno-3-phosphoglycerate does not engage in the next glycolytic reaction, which normally produces one ATP molecule, but it falls back to arsenate and 3-phosphoglycerate, without the production of ATP, while the arsenate released can act again on the enzyme in a cyclical manner. The result is that the glycolytic pathway is uncoupled from ATP-production. Needless to say this signifies a severe and often fatal inhibition of energy metabolism.

Species patterns of metal susceptibility

Animals, plants, fungi, protists and prokaryotes all differ greatly in their susceptibility to metals. To give a few examples:

- Earthworms and snails are known to be quite sensitive to copper; the absence of earthworms in orchards and vineyards where copper-containing fungicides are used is well documented. Snails cannot be cultured in water that runs through copper-containing linings. Fungi are also sensitive to copper, which explains the use of copper in fungicides. Also many plants are sensitive to copper due to the effects on root growth. Among vertebrates, sheep are quite sensitive to copper, unlike most other mammals.
- Crustaceans as well as fish are relatively sensitive to zinc. Mammals, however, are hardly sensitive to zinc at all.
- Humans are relatively sensitive to lead because high lead exposure lead disturbs the development of children's brain and is correlated with low IQ-scores. Most invertebrates however, are quite insensitive to lead.
- Although many invertebrates are quite sensitive to cadmium, the interspecies variation in sensitivity to this element is particularly high, even within the same phylogenetic lineage. The soil-living oribatid mite *Platynothrus peltifer* is one of the most sensitive invertebrates with respect to the effect of cadmium on reproduction, however *Oppia nitens*, also an oribatid, is extremely tolerant to cadmium.

In the end, such patterns must be explained in terms of the presence of susceptible biochemical targets, different strategies for storage and excretion, and differing mechanisms of defence and sequestration. However, at the moment there is no general framework by which to compare the variation of sensitivity across species. Also, there is no relation between accumulation and susceptibility; some species that accumulate metals to a large degree (e.g. copper in isopods) are not sensitive to the same metal, while others, which do not accumulate the metal, are quite sensitive. Accumulation seems to be partly related to a species feeding strategy (e.g. spiders absorb almost al the (fluid) food they take in and any metals in the food will accumulate in the midgut gland); accumulation is also related to specific nutrient requirements (e.g. copper in isopods, manganese in some oribatid mites). Finally, some populations of some species have evolved specific tolerances in response to their living in a metal-contaminated environment, on top of the already existing accumulation and detoxification strategies.

Conclusion

Metals do not form a homogeneous group. Their toxicity involves reactivity towards a great variety of biochemical targets. Often several mechanisms act simultaneously and interact with each other. Induction of oxidative stress is a common denominator, as is reaction to nucleophilic groups in macromolecules. The great variety of metal-induced responses makes them interesting model compounds for toxicological studies.

References

Cameron, K.S., Buchner, V., Tchounwou, P.B. (2011). Exploring the molecular mechanisms of nickel-induced genotoxicity and carcinogenicity: a literature review. Reviews of Environmental Health 26, 81-92.

Ernst, W.H.O., Joosse-van Damme, E.N.G. (1983) *Umwelbelastung durch Mineralstoffe*. Fischer Verlag, Jena.

Singh, A.P., Goel, R.K., Kaur, T. (2011) Mechanisms pertaining to arsenic toxicity. Toxicology International 18, 87-93.

Verbost, P.M. (1989) *Cadmium toxicity: interaction of cadmium with cellular calcium transport mechanisms* Ph.D. thesis, Radboud Universiteit Nijmegen.

4.2.11. Question 1

Mention three different classes of primary lesions due to free metal ions and causing metal toxicity. Include the metals that are best known for causing each type of lesion.

4.2.11. Question 2

Is it possible to decide to which type of metal a cell is exposed, based on the kind of cellular disturbance that is observed?

4.2.11. Question 3

Several invertebrate accumulate metals to a very high degree. Mention a few examples and the metals they accumulate. Are these animals also among the most sensitive to metal toxicity? Please explain.

4.2.11. Question 4

"Essential metals can be regulated and are therefore less toxic than xenobiotic metals" - Please comment on this thesis.

4.2.12. Metal tolerance

Author: Nico M. van Straalen

Reviewers: Henk Schat, Jaco Vangronsveld

Learning objectives

You should be able to

- describe which mechanisms of changes in metal trafficking can contribute to metal tolerance and hyperaccumulation
- explain the molecular factors associated with the evolution of metal tolerance in plants and in animals
- develop an opinion on the issue of "rescue from pollution by evolution" in the risk assessment of heavy metals

Keywords: hyperaccumulation, metal uptake mechanisms, microevolution

Synopsis

Some species of plants and animals have evolved metal-tolerant populations that can survive exposures that are lethal for other populations of the same species. Best known is the heavy metal vegetation that grows on metalliferous soils. The study of these cases of "evolution in action" has revealed many aspects of metal trafficking in plants, transport across membranes, metal scavenging molecules in the cell, and subcellular distribution of metals, and how these processes have been adapted by natural selection for tolerance. Metal-tolerant plant varieties are usually dependent upon high metal concentrations in the soil and do not grow well in reference soils. In addition, some plant species show an extreme degree of metal accumulation. In animals metal tolerance has been demonstrated in some invertebrates that live in close contact with metal-containing soils and this is usually achieved by altered regulation of metal scavenging proteins such as metallothioneins, or by duplication of the corresponding genes. Genomics studies are broadening our perspective as the adaptation normally does not rely on a single gene but includes hypostatic factors and modifiers.

Introduction

As metals cannot be degraded or metabolized, the only way to deal with potentially toxic excess is to store or excrete them. Often both mechanisms are operational, excretion being preceded by storage or scavenging, but animals and plants differ greatly in the emphasis on one or the other mechanism. Both essential and nonessential metals are subject to all kinds of trafficking mechanisms aiming to keep the biologically active, free ion concentration of the metal extremely low. Still, there is hardly any relationship between accumulation and tolerance. Some species have low tissue concentrations and are sensitive, others have low tissue concentrations and are tolerant, some accumulate metals and suffer from the high concentrations, others accumulate and are extremely tolerant.

Like the mechanisms of biotransformation (see the section on Genetic Variation) metal trafficking mechanisms show genetic variation and such variation may be subject to evolution. However, it has to be noted that only in a limited number of plants and animal species metal-tolerant populations have evolved. This may be due to the fact that evolution of metal tolerance makes use of already existing, moderately efficient, metal trafficking mechanisms in the ancestral species. This interpretation is suggested by the observation that the non-metal-tolerant varieties of metal-tolerant plants already have a certain degree of metal tolerance (larger than species that never evolve metal-tolerant varieties). So the mutational distance to metal tolerance was smaller in the ancestors of metal-tolerant plants than it is in "normal" plants.

Real metal tolerance, where the metal-tolerant population can withstand orders of magnitude larger exposures than reference populations, and has become dependent on metal-rich soils, is only found in plants. Metal tolerance in animals is of degree, rather than of kind, and does not come with externally recognizable phenotypes. Most likely the combination of strong selection pressure, the impossibility to escape by locomotion and the right pre-existing genetic variation explain why metal tolerance in plants is so much more prominent compared to animals.

In this section we will discuss the various mechanisms that have been shown to underlie metal tolerance. The evolutionary response to environmental metal exposure is one of the classical examples of "evolution in action", next to insecticide resistance in mosquitoes and industrial melanism in butterflies.

Metal tolerance in plants

For many years, most likely already since humans started to dig ores and use metals for the manufacture of utensils, pottery and tools, it has been known that naturally metal-rich soils harbour a specific metal-tolerant vegetation. This "Schwermetallvegetation", described in the classical book by the German-Dutch botanist W.H.O. Ernst, consists of a designated collection of plant species, with representatives from various families. Several species also have metal-sensitive populations living in normal soils, but some, like the European zinc violet, *Viola calaminaria*, are restricted to metal-rich soils. This is also seen in the metal-tolerant vegetations of New Caledonia, Cuba, Zimbabwe and Congo, which to a large degree consist of endemic metal-tolerant species (true metallophytes) that are never found in normal soils. However, some common species also developed metal-tolerant ecotypes.

Metal-tolerant plant species have expanded their range when humans started to dig the metal ores and now can also be found extensively at mining sites, metal-enriched stream banks, and around metal smelters. Naturally metal-enriched soils differ from reference soils not only in metal concentration but also in other aspects, e.g. calcium and moisture, so the selection for metal tolerance comes goes hand-in-hand with selection by several other factors.

Metal tolerance is mainly restricted to herbs and forbs, and (except some tropical serpentines) does not extend to trees. A heavy metal vegetation is recognizable in the landscape as a "meadow", lacking trees, with relatively few plant species and an abundance of metallophytes. In the past, metal ores were discovered from the presence of such metallophytes, an activity called bioprospecting.

We know from biochemistry that different metals are bound to different ligands and follow different biochemical pathways in biological tissues (see the section on metal accumulation). Some metals (cadmium, copper, mercury) are "sulphur-seekers", others have an affinity to organic acids (zinc) and still others tend to be associated with calcium-rich tissues (lead). Essential metals such as copper, zinc and iron have their own, metal-specific, transport mechanisms. From these observations one may conclude that metal tolerance will also be specific to the metal and that cross-tolerance (tolerance directed to one metal causing tolerance to another metal as a side-effect) is relatively rare. This is indeed the case.

In many cases metal-tolerant plants do not show the same growth characteristics as the non-tolerant varieties of the same species. Loss of growth potential has often been interpreted as a "cost of tolerance". However, genetic research has shown that the lower growth potential of metallophytes is a separate adaptation, to deal with the usually infertile metalliferous soils, and there is no mechanistic link to tolerance. Metabolic costs or negative pleiotropic effects of metal tolerance have not been described. The fact that metal-tolerant plants do not grow well in clean soils is explained by the constitutive upregulation of trafficking and compartmentalization mechanisms, causing increased metal requirements that cannot be met on non-metalliferous soils.

Another striking fact is that metal tolerances in the same plant species at different sites have evolved independently from each other. The various metal-tolerant populations of a species do not all descend from a single ancestral population, but result from repeated local evolution. That still in different populations sometimes the same loci are affected by natural selection, is ascribed to the fact that, given the species' genetic background, there are only a limited number of avenues to metal tolerance.

A final general principle is that metal tolerance in plants is often targeted towards proteins that transport metals across membranes (cell membrane, tonoplast). The genes of such transporters may be duplicated, the balance between high-affinity transporters and low-affinity versions may be altered, their expression may be upregulated or downregulated, or the proteins may be targeted to different cellular compartments.

Although many details on the genetic changes responsible for tolerance in plants are still lacking, the work on copper tolerance in bladder campion, *Silene vulgaris*, illustrates many of the points listed above. The plant has many metal-tolerant populations, of which one found at Imsbach, Germany, shows an extreme degree of copper tolerance and also some (independently evolved) zinc and cadmium tolerance. The area is known for its "Bergbau" with historical mining activities for copper, silver and cobalt, but also some older calamine deposits, which explains the zinc and cadmium tolerance.

Genetic work by H. Schat and colleagues has shown that two ATP-driven copper transporters, designated *HMA5I* and *HMA5II* are involved in copper tolerance of *Silene*. The HMA5I protein resides in the tonoplast to relocate copper into the vacuole, while HMA5II resides in the endoplasmic reticulum. When free copper ions appear in the cell, HMA5II relocates from the ER to the cell membrane and starts pumping copper out of the cell. During transport from roots to shoot (in the xylem vessels) copper is bound as a nicotianamine complex. In addition, plant metallothioneins play a role in copper binding and transport in the phloem and during redistribution from senescent leaves. Copper tolerance in *Silene* illustrates the principle referred to above that metal tolerance is achieved by enhancing the transport mechanisms already present, not by evolving new genes.

Metal hyperaccumulation

Some plants accumulate metals to an extreme degree. Well-known are metallophytes growing on serpentine soils, which accumulate very large amounts of nickel. Also copper and cobalt accumulation is observed in several species of plants. Hyperaccumulators do not exclude metals but preferentially accumulate them when the concentration in the soil is extremely high (> 50.000 mg of copper per kg soil). The copper concentration of the leaves may reach values of more than 1000 µg/g. A very extreme example is a tree species, *Sebertia acuminata*, growing on the island of New Caledonia in ultramafic soil with 0.85% of nickel, which produces a latex containing 11% of nickel by weight. Such extraordinary high concentrations impose extreme demands on the efficiency of metal trafficking and so have attracted the attention of biological investigators. In Western Europe's heavy metal vegetation, zinc accumulators are present in several species of the genera *Agrostis*, *Brassica*, *Thlaspi* and *Silene*.

Most of the experimental research is conducted on the brassicacean species *Noccaea (Thlaspi) caerulescens* and *Arabidopsis halleri*, with *Arabidopsis thaliana* as a non-accumulating reference model.

The transport of metals in a plant involves a number of distinct steps, where each step is upregulated in the metal hyperaccumulator. This illustrated in Figure 1 in Verbruggen et al. (2009) for zinc hyperaccumulation in *Thlaspi caerulescens*.

- Uptake in root epithelial cells; this involves ZIP4 and IRT1 zinc transporters
- Transport between root tissues
- Loading of the root xylem, by means of HMA4 and other metal transporters
- In the xylem zinc may be chelated by citrate, histidine of nicotianamine, or may just be present as free ions
- Unloading of the xylem in the leaves. This involves YSL proteins and others.
- Transport into vacuoles and chelation to vacuole-specific chelators such as malate, involving metal transporters such as HMA3, MTP1 and MHX.

While the basic components of the system are beginning to be known, the question how the whole machinery is upregulated in a coherent fashion is not yet clear.

Metal tolerance in animals

Also in animals, metal tolerant populations of the same species have been reported, however, there is no specific metal-tolerant community with a designated set of species, like in plants. There are, however, obvious metal accumulators among animals. Best known are terrestrial isopods, which accumulate very high concentrations of copper in designated cells in the their hepatopancreas, and some species of oribatid mites which accumulate very high amounts of manganese and zinc.

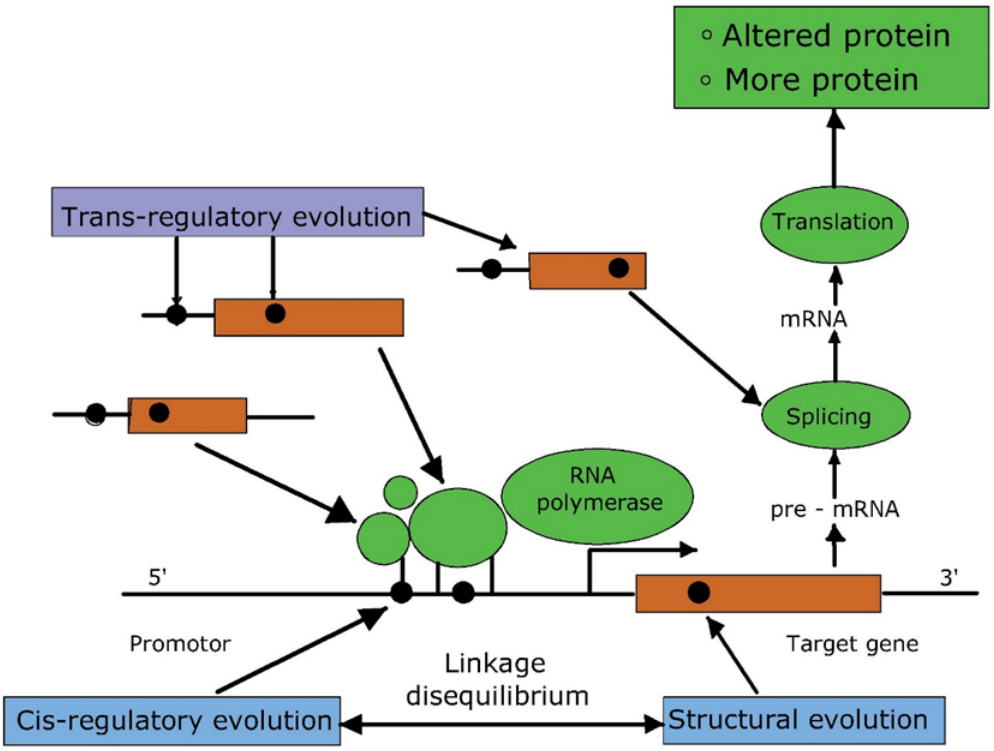


Figure 2. Simplified scheme of transcriptional regulation of a gene involved in metal detoxification, such as metallothionein. The dots indicate the various mutations possible. Two different pathways to tolerance are sketched: structural mutations altering the protein (e.g. increasing binding affinity) and regulatory mutations altering the amount of protein by regulating transcription. Transcriptional regulation can be in cis (changes in the promoter, affecting the binding of transcription factors) or in trans (transcription factor or other regulatory proteins). Redrawn from Van Straalen et al. (2011) by Wilma IJzerman.

One of the factors investigated to explain metal tolerance in animals is a metal-binding protein, metallothionein (MT). Gene duplication of an *MT* gene has been implicated in the tolerance of *Daphnia* and *Drosophila* to copper. In addition, metal tolerance may be due to altered transcriptional regulation. The latter mechanism underlies the evolution of cadmium tolerance in the soil-living springtail, *Orchesella cincta*. Detailed genetic analysis of this model system has revealed that the MT promoter of *O. cincta* shows a very large degree of polymorphism, some alleles affecting the transcription factor binding sites and causing overexpression of MT. The promoter allele conferring strong overexpression of MT upon exposure to cadmium, had a significantly higher frequency in *O. cincta* populations from metal-contaminated soils (Figure 2).

In addition to springtails, evolution of metal tolerance has also been described for the earthworm, *Lumbricus rubellus*. In a population living in a lead-contaminated deserted mining area in Wales two lineages were distinguished on the basis of the COI gene and RFLPs, Interestingly, the two lineages had colonized different microhabitats of the area, one of them being unable to survive high lead concentrations. Differential expressions were noted for genes in phosphate and calcium metabolism. Two crucial mutations in a calcium transport protein suggested that lead tolerance in *L. rubellus* is due to modification of calcium transport, a logical target since lead and calcium are often found to interact with each other's transport (see the section on metal accumulation).

Conclusions

The study of metal tolerance is a rewarding topic of evolutionary ecotoxicology. Several crucial genetic mechanisms have been identified but in none of the study systems a complete picture of the evolved tolerance mechanisms is available. It may be expected that genome-wide studies will be able to identify the full network responsible for tolerance, which most likely includes not only major genes, but also hypostatic factors and modifiers.

References

Andre, J., King, R.A., Stürzenbaum, S.R., Kille, P., Hodson, M.E., Morgan, A.J. (2010). Molecular genetic differentiation in earthworms inhabiting a heterogeneous Pb-polluted landscape. Environmental Pollution 158, 883-890.

Ernst, W.H.O. (1974) *Schwermetallvegetation der Erde*. Gustav Fischer Verlag, Stuttgart.

Janssens, T.K.S., Roelofs, D., Van Straalen, N.M. (2009). Molecular mechanisms of heavy metal tolerance and evolution in invertebrates. Insect Science 16, 3-18.

Krämer, U. (2010). Metal hyperaccumulation in plants. Annual Review of Plant Biology 61, 517-534.

Li, X., Iqbal, M., Zhang, Q., Spelt, C., Bliek, M., Hakvoort, H.W.J., Quatrocchio, F.M., Koes, R., Schat, H. (2017). Two *Silene vulgaris* copper transporters residing in different cellular compartments confer copper hypertolerance by distinct mechanisms when expressed in *Arabidopsis thaliana*. New Phytologist 215, 1102-1114.

Lopes, I., Baird, D.J., Ribeiro, R. (2005). Genetically determined resistance to lethal levels of copper by *Daphnia longispina*: association with sublethal response and multiple/coresistance. Environmental Toxicology and Chemistry 24, 1414-1419.

Van Straalen, N.M., Janssens, T.K.S., Roelofs, D. (2011). Micro-evolution of toxicant tolerance: from single genes to the genome's tangled bank. Ecotoxicology 20, 574-579.

Verbruggen, N., Hermans, C., Schat, H. (2009) Molecular mechanisms of metal hyperaccumulation in plants. New Phytologist 181, 759-776.

4.2.12. Question 1

Describe the pathway of zinc ions taken up by plants from the soil solution to their final destination in the plant, and how the various steps have been modified in hyperaccumulation plants species such as *Arabidopsis halleri*.

4.2.12. Question 2

Discuss the difference between trans-regulatory change and cis-regulatory change in the evolution of metal tolerance.

4.2.13. Adverse Outcome Pathways

Author: Dick Roelofs

Reviewers: Nico van Straalen, Dries Knapen

Learning objectives:

You should be able to

- explain the concept of adverse outcome pathway.
- interpret a graphical representation of an AOP
- search the AOPwiki database for molecular initiating events, key events and adverse outcomes

Keywords: Molecular initiation event, key event, in vitro assay, high throughput assay, pathway

Introduction

Over the past two decades the availability of molecular, biochemical and genomics data has exponentially increased. Data are now available for a phylogenetically broad range of living organisms, from prokaryotes to humans. This has tremendously advanced our knowledge and mechanistic understanding of biological systems, which is highly beneficial for different fields of biological research such as genetics, evolutionary biology and agricultural sciences. Being an applied biological science, toxicology has not yet tapped this wealth of data, because it is difficult to incorporate mechanistic data when assessing chemical safety in relation to human health and the environment. However, society is increasingly concerned about the release of industrial chemicals with little or no hazard- or risk information. Consequently, a much larger number of chemicals need to be considered for potential adverse effects on human health and ecosystem functioning. To meet this challenge it is necessary to deploy fast, cost-effective and high throughput approaches that can predict potential toxicity of substances and replace traditional tests based on survival and reproduction that run for weeks or months and often are quite labour-intensive. A major challenge is however, to link these fast *in vitro* and *in vivo* assays to endpoints used in current risk assessment. This challenge was picked up by defining the adverse outcome pathway (AOP) framework, for the first time proposed by the Gerald Ankley and co-workers from the United States Environmental Protection Agency, US-EPA (Ankley et al., 2010).

The framework

The AOP framework is defined as an evolution of prior pathway-based concepts, most notably mechanisms and modes of action, for assembling and depicting toxicological data across biological levels of organization (Ankley and Edwards, 2018). An AOP is a graphical representation of a series of measurable *key events* (KEs). A key event is a measurable directional change in the state of a biological process. KEs can be linked to one another through *key event relationships* (KERs; see Figure 1). The first KE is depicted as the "*molecular initiating event*" (MIE), and represents the interaction of the chemical with a biological receptor that activates subsequent key events. The key event relationships should ideally be based on causal evidence. A cascade of key events can eventually result in an *adverse outcome* (AO) at the individual or population level. The MIE and AO are specialized KEs, but treated like any other KE in the AOP framework.

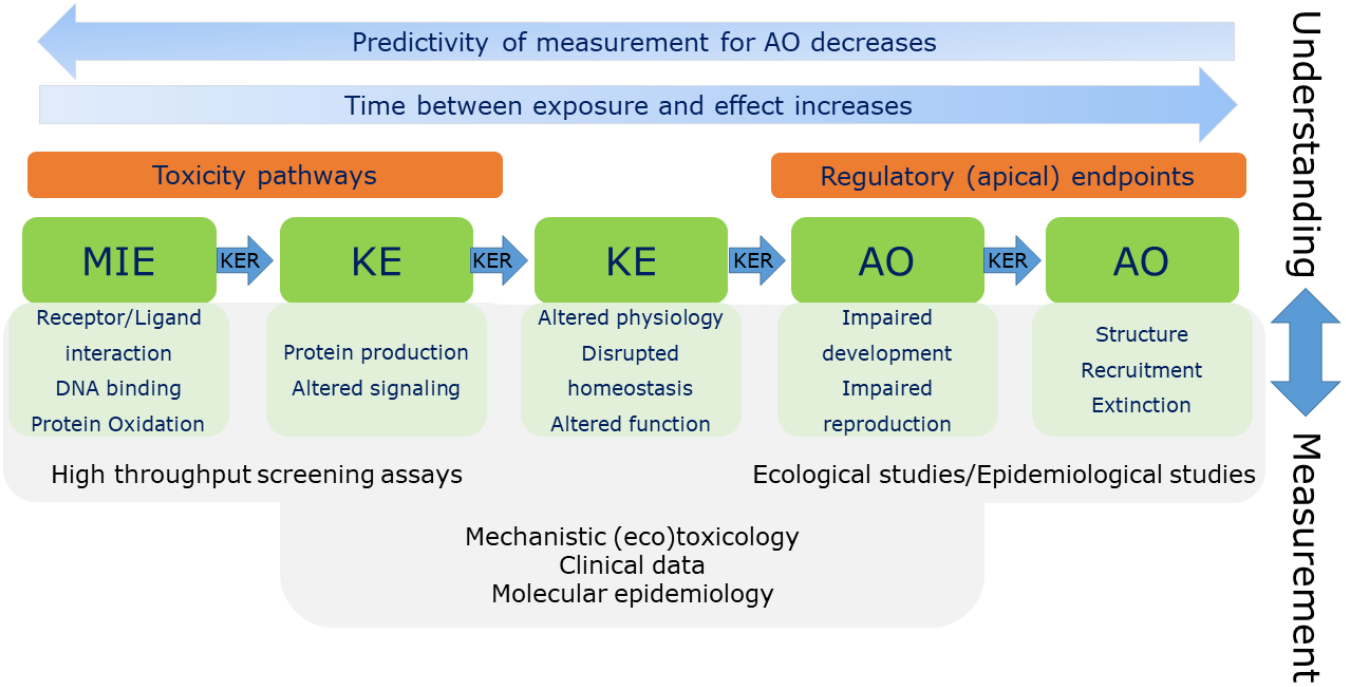


Figure 1. The Adverse Outcome Pathway framework. Various measurable data streams are linked in a causal manner and eventually are related to outcomes essential for risk assessment. MIE, molecular initiating event; KE, key event; KER, key event relationship; AO, adverse outcome. Adapted from Ankley and Edwards (2018) by Kees van Gestel.

The aim of an AOP is to represent and describe, in a simplified way, how responses at the molecular- and cellular level are translated to impacts on development, reproduction and survival, which are relevant endpoints in risk assessment (Villeneuve et al., 2014). Five core concepts have been defined in the development of AOPs:

1. AOPs are not chemical specific, they are biological pathways;
2. AOPs are modular, they refer to a designated and defined metabolic cascade, even if that cascade interacts with other biological processes;
3. individual AOPs are developed as pragmatic units;
4. networks of multiple AOPs sharing KEs and KERs are functional units of prediction for real-world scenarios; and
5. AOPs are living documents that may change over time based on new scientific insights.

Generally, AOPs are simplified linear pathways but different AOPs can be organized in networks with shared nodes. The AOP networks are actually the functional units of prediction, because they represent the complex biological interactions that occur in response to exposure to a toxicant or a mixture of toxicants. Analysis of the intersections (shared key events) of different AOPs making up a network can reveal unexpected biological connections (Villeneuve et al., 2014).

Molecular initiating events and key events

Typically, an AOP consists of only one MIE, and one AO, connected to each other by a potentially unlimited number of KEs and KERs. The MIE is considered to be the first anchor of an AOP at the molecular level, where stressors directly interact with the biological receptor. Identification of the MIE mostly relies on chemical analysis, *in silico* analysis or *in chemico* and *in vitro* data. For instance, the MIE for AOPs related to estrogen receptor activation involves the binding of chemicals to the estrogen receptor, thereby triggering a cascade of effects in hormone-related metabolism (see the section on Endocrine disruption). The MIE for AOPs related to skin sensitization (see below) involves the covalent interaction of chemicals to skin proteins in skin cells, an event called haptenization (Vinken, 2013).

A wide range of biological data can support the understanding of KEs. Usually, early KEs (directly linked to MIEs) are assessed using *in vitro* assays, but may include *in vivo* data at the cellular level, while intermediate and late KEs rely on tissue-, organ- or whole organism measurements (Figure 1). Key-event measurements are also related to data from high-throughput screening and/or data generated by different -omics technologies. This is actually where the true value of the AOP framework comes in, since it is currently the only framework able to reach such a high level of data integration in the context of risk assessment. It is even possible to integrate comparative data from phylogenetically divergent organisms into key event measurements, valid across species, which could facilitate the evaluation of species sensitivity (Lalone et al., 2018). The final AO output is usually represented by apical responses, already described as standard guidelines accepted and instrumental in regulatory decision-making, which include endpoints such as development, growth, reproduction and survival.

Development of the AOP framework is currently supported by US-Environmental Protection Agency, EU Joint Research Centers (ERC) and the Organization for Economic Cooperation and Development (OECD). Moreover, OECD has sponsored the development of an open access searchable database AOPWiki (https://aopwiki.org/), comprising over 250 AOPs with associated MIEs, KEs and KERs, and more than 400 stressors. New AOPs are added regularly. The database also has a system for specifying the confidence to be placed in an AOP. Where KEs and KERs are supported by direct, specifically designed, experimental evidence, high confidence is placed in them. In other cases confidence is considered moderate or low, e.g. when there is a lack of supporting data or conflicting evidence.

Case example: covalent protein binding leading to skin sensitization (AOP40)

Skin sensitization is characterized by a two-step process, a sensitization phase and an elicitation phase. The first contact of electrophile compounds with the skin covalently modifies skin proteins and generates an immunological memory due to generated antigen/allergen specific T-cells. During the elicitation phase, repeated contact with the compound elicits the allergic reaction defined as allergic contact dermatitis, which usually develops into a lifelong effect. This is an important endpoint for safety assessment of personal care products, traditionally evaluated by *in vivo* assays. Based on changed public opinion the European Chemical Agency (ECHA) decided to move away from whole animal skin tests, and developed alternative assessment strategies. During sensitization, the MIE takes place when the chemical enters the skin, where it forms a stable complex with skin-specific carrier proteins (hapten complexes), which are immunogenic. A subsequent KE comprises inflammation and oxidative defense via a signaling cascade called the Keap1/Nrf2 signalling pathway (*Kelch-like ECH-associated protein 1 / nuclear factor erythroid 2 related factor 2*). At the same time, a second KE is defined as dendritic cell activation and maturation. This results into movement of dendritic cells to lymph nodes, where the hapten complex is presented to naive T-cells. The third KE describes the proliferation of hapten-specific T-cells and subsequent movement of antigen-specific memory cells that circulate in the body. Upon a second contact with the compound, these memory T-cells secrete cytokines that cause an inflammation reaction leading to the AO including red rash, blisters, and burning skin (Vinken et al., 2017). This AOP is designated AOP40 in the database of adverse outcome pathways.



Figure 2. Adverse Outcome Pathway for covalent protein binding leading to skin sensitization (AOP40). Adapted from Vinken et al. (2017) by Kees van Gestel.

A suite of high throughput *in vitro* assays have now been developed to quantify the intermediate KEs in AOP40. These data formed the basis for the development of a Bayesian network analysis that can predict the potential for skin sensitization. This example highlights the use of pathway-derived data organized in an AOP, ultimately leading to an alternative fast screening method that may replace a conventional method using animal experiments.

References

Ankley, G.T., Bennett, R.S., Erickson, R.J., Hoff, D.J., Hornung, M.W., Johnson, R.D., Mount, D.R., Nichols, J.W., Russom, C.L., Schmieder, P.K., Serrrano, J.A., Tietge, J.E., Villeneuve, D.L. (2010). Adverse outcome pathways: A conceptual framework to

support ecotoxicology research and risk assessment. Environmental Toxicology and Chemistry 29, 730-741.

Ankley, G.T., Edwards, S.W. (2018). The adverse outcome pathway: A multifaceted framework supporting 21st century toxicology. Current Opinion in Toxicology 9, 1-7.

LaLone, C.A., Villeneuve, D.L., Doering, J.A., Blackwell, B.R., Transue, T.R., Simmons, C.W., Swintek, J., Degitz, S.J., Williams, A.J., Ankley, G.T. (2018). Evidence for cross species extrapolation of mammalian-based high-throughput screening assay results. Environmental Science and Technology 18, 13960-13971.

Villeneuve, D.L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T.H., LaLone, C.A., Landesmann, B, Lettieri, T., Munn, S., Nepelska, M., Ottinger, M.A., Vergauwen, L., Whelan, M. (2014). Adverse Utcome Pathway development I: Strategies and principles. Toxicological Sciences 142, 312-320.

Vinken, M. (2013). The adverse outcome pathway concept: A pragmatic tool in toxicology. Toxicology 312, 158-165.

Vinken, M., Knapen, D., Vergauwen, L., Hengstler, J.G., Angrish, M., Whelan, M. (2017). Adverse outcome pathways: a concise introduction for toxicologists. Archives of Toxicology 91, 3697-3707.

4.2.13. Question 1

What is a molecular initiation event?

4.2.13. Question 2

Where in the chain of events do high-throughput *in-vitro* assays feed into AOPs?

4.2.13. Question 3

What is the ultimate goal of the AOP concept?

4.2.13. Question 4

How many intermediate key events are captured in the AOP for skin sensitization?

4.2.13. Question 5

Is an AOP always represented as a linear chain between MIE and AO?

4.2.14. Genetic variation in toxicant metabolism

Author: Nico M van Straalen

Reviewers: Andrew Whitehead, Frank van Belleghem

Learning objectives:

You should be able to

- explain four different classes of CYP gene variation and expression, contributing to species differences
- explain the associations between biotransformation activity and specific ecologies
- explain how genetic variation in biotransformation enzymes may lead to evolution of toxicant tolerance
- describe the relevance of human genetic polymorphisms for personalized medicine

Keywords: toxicant susceptibility; genetic variation; biotransformation evolution of toxicant tolerance

Assumed prior knowledge and related modules

- Biotransformation and internal processing of chemicals
- Defence mechanisms
- Genetic erosion

In addition, a basic knowledge of genetics and evolutionary biology is needed to understand this module.

Synopsis

Susceptibility to toxicants often shows inter-individual differences associated with genetic variation. While such differences are considered a nuisance in laboratory toxicity testing, they are an inextricable aspect of toxicant effects in the environment. Variation may be due to polymorphisms in the target site of toxicant action, but more often differences in metabolic enzymes and rates of excretion contribute to inter-individual variation. The structure of genes encoding metabolic enzymes, as well as polymorphisms in promoter regions of such genes are common sources of genetic variation. Under strong selection pressure species may evolve toxicant-tolerant populations, for example insects to insecticides and bacteria to antibiotics. In human populations, polymorphisms in drug metabolizing enzymes are mapped to provide a basis for personal therapies. This module aims to illustrate some of the genetic principles explaining inter-individual variation of toxicant susceptibility and its evolutionary consequences.

Introduction

For a long time it has been known that human subjects may differ markedly in their responses to drugs: while some patients hardly respond to a certain dosage, others react vehemently. Similar differences exist between the sexes and between ethnic groups. To avoid failure of treatment on the one hand and overdosing on the other, such personal differences have attracted the interest of pharmacological scientists. Also the tendency to develop cancer upon exposure to mutagenic chemicals is partly due to genetics. Since the rise of molecular ecology in the 1990s ecotoxicologists have noted that inter-individual differences in toxicant responses also exists in the environment.

Due to genetic variation environmental pollution may trigger evolutionary change in the wild. From quantitative genetics we know that when a trait is due to many genes, each with an independent additive effect on the trait value, the response to selection R, is linearly related to the selection differential S according to the formula: $R = h^2S$, where h^2 is a measure of the heritability of the selected trait (fraction of additive genetic variance relative to total phenotypic variance). Since anthropogenic toxicants can act as very strong selective agents (large S) it is expected that whenever $h^2 > 0$ there will be adaptation. However, the effectiveness of "evolutionary rescue" from pollution is limited to those species that have the appropriate genetic variation and the ability to quickly increase in population size.

Polymorphisms of drug metabolizing enzymes in humans

One of the most important enzyme systems contributing to metabolism of xenobiotic chemicals is the cytochrome P450 family, a class of proteins located in the smooth endoplasmic reticulum of the cell and acting in co-operation with several other proteins. Cytochrome P450 will oxidize the substrate and enhance its water-solubility (called phase-I reaction), and in many cases activate it for further reactions involving conjugation with an endogenous compound (phase II reactions). These processes generally lead to detoxification and increased excretion of toxic substances. The biochemistry of drug metabolism is discussed in detail in the section on Xenobiotic metabolism and defence.

The human genome has 57 genes encoding a P450 protein. The genes are commonly designated as "CYP". Other organisms, especially insects and plants have many more CYPs. For example, the *Drosophila* genome encodes 83 functional P450 genes and the genome of the model plant *Arabidopsis* has 244 CYPs. Based on sequence similarity, CYPs are classified in 18 families and 43 subfamilies, but there is no agreement yet about the position of various CYP genes in lower invertebrates. The complexity is enhanced by duplications specific to certain evolutionary lineages, creating a complicated pattern of orthologs (homologs by descent from a common ancestor) and paralogs (homologs due to duplication in the same genome). In addition to functional enzymes it is also common to find many CYP pseudogenes in a genome. Pseudogenes are DNA-sequences that resemble functional genes, but are mutated and they do not result in functional proteins).

The expression of CYP enzymes is markedly tissue-specific. Often CYP expression is high in epithelial tissues (lung, intestine) and organs with designated metabolic activity (liver, kidney). In the human body, the liver is the main metabolic organ and is known for its extensive CYP expression. P450 enzymes also differ in their inducibility by classes of chemicals and in their substrate specificity.

It is often assumed that the versatility of an organism's CYP genes is a reflection of its ecology. For example, herbivorous insects that consume plants of different kinds with many different feeding repellents must avail of a wide diversity of CYP genes. It has also been shown that activity of CYP enzymes among terrestrial organisms is, in general, higher than among aquatic organisms and that plant-eating birds have higher biotransformation activities than predatory birds.

One of the best-investigated CYP genes, especially due to its strong inducibility and involvement in xenobiotic metabolism, is mammalian CYP1A1. In humans induction of this gene is associated with increased lung cancer risk from smoking, and with other cancers, such as breast cancer and prostrate cancer. Human CYP1A1 is located on chromosome 15 and encodes 251 amino acids in seven exons (see Figure 2 in Zhou et al., 2009). About 133 single-nucleotide polymorphisms (SNPs, variations in a single nucleotide that occur at a specific position in the genome) have been described for this gene, of which 23 are non-synonymous (causing a substitution of an amino acid in the protein).

Many of these SNPs have a medical relevance. For example, a rather common SNP in exon 7 changes codon 462 from isoleucine into valine. The substituted allele is called CYP1A1*2A, and this occurs at a frequency of 19% in the Caucasian part of the human population. The allelic variant of the enzyme has a higher activity towards 17β-estradiol and is a risk factor for several types of cancer. However, the expression of such traits may vary from one population to another, and may also interact with other risk factors. For example, CYP1A1*2A is a risk factor for cervical cancer in women with a history of smoking in the Polish population, but the same SNP may not be a risk factor in another population or among people with a non-smoking lifestyle. In genetics these effects are known as epistasis: the phenotypic effect of genetic variation at one locus depends on the genotype of another locus. This is also an example of a genotype-by-environment interaction, where the phenotypic effect of a genetic variant depends on the environment (smoking habit). In toxicology it is known that polymorphisms of phase II biotransformation enzymes may significantly contribute to epistatic interaction with CYP genes. Unraveling all these complicated interactions is a very active field of research in human medical genetics.

Cytochrome P450 variation across species

Comparison of CYP genes in different species has revealed an enormously rapid evolution of this gene family, with many lineage-specific duplications. This indicates strong selective pressures imposed by the need to detoxify substances ingested with the diet. Especially herbivorous animals are constantly exposed to such compounds, synthesized by plants to deter feeding. We also see profound changes in CYP genes associated with evolutionary transitions such as colonization of terrestrial habitats by the various lineages of arthropods. Such natural variation, induced by plant toxins and habitat requirements, is also relevant in the responses to toxicants.

In general, variation of biotransformation enzymes can be classified in four main categories

1. Variation in the structure of the genes, e.g. substitutions that alter the binding affinity to substrates; such variation discriminates the various CYP genes.
2. Copy number variation; duplication usually leads to an increase in enzymatic capacity; this process has been enormously important in CYP evolution. Because CYP gene duplications are often specific to the evolutionary lineage, a complicated pattern of paralogs (duplicates within the same genome) and orthologs (genes common by descent, shared with other species) arises.
3. Promoter variation, e.g. due to insertion of transposons or changes in the number or arrangement of transcription factor binding sites. This changes the amount of protein produced from one gene copy by altered transcriptional regulation.
4. Variation in the structure, action or activation of transcriptional regulators. The transcription of biotransformation enzymes is usually induced by a signaling pathway activated by the compound to be metabolized (see the section on Xenobiotic metabolism and defence), and this pathway may show genetic variation.

To illustrate the complicated evolution of biotransformation genes, we shortly discuss the CYPs of common cormorant, *Phalacrocorax carbo*. This is a bird known for its narrow diet (fish) and extraordinary potential for accumulation of dioxin-related compounds (PCBs, PCDDs and PCDFs). Environmental toxicologists have identified two CYP1A genes in the cormorant, called CYP1A4 and CYP1A5. It turns out that CYP1A4 is homologous by descent (orthologous) to mammalian CYP1A1 while CYP1A5 is an ortholog of mammalian CYP1A2. However, the orthologies are not revealed by common phylogenetic analysis if the whole coding sequence is used in the alignment (see Figure 3 in Kubota et al., 2006). This is a consequence of a process called interparalog gene conversion, which tends to homogenize DNA sequences of gene copies located on the same chromosome. This diminishes sequence variation between the paralogs, and creates chimeric gene structures, that are more similar to each other than expected from their phylogenetic relations. If a phylogenetic tree is made using a section of the gene that remained outside the gene conversion, the true phylogenetic relations are revealed (see Figure 3 in Kubota et al., 2006).

Cytochrome P450-mediated resistances

Cytochrome P450 polymorphisms are also implicated in certain types of insecticide resistance. There are many ways in which insects and other arthropods can become resistant and several mechanisms may even be present in the same resistant strain. Target site alteration (making the target less susceptible to the insecticide, e.g. altered acetylcholinesterase, substitutions in the GABA-receptor, etc.) seems to be the most likely mechanism for resistance, however, such changes often come with substantial costs as they may diminish the natural function of the target (in genetics this is called pleiotropy). Increased metabolism does not usually contribute metabolic costs and this is where cytochromes P450 come into play. A model system for investigating the genetics of such mechanisms is DDT resistance in the fruit fly, *Drosophila melanogaster*.

In a DDT-resistant *Drosophila* strain, all CYP genes were screened for enhanced expression and it was shown that DDT resistance was due to a highly upregulated variant of only a single gene, Cyp6g1. Further analysis showed that the gene's promoter carried an insertion with strong similarity to a transposable element of the *Accord* family. The insertion of this element causes a significant overexpression and a high rate of protein synthesis that allows the fly to quickly degrade a DDT dose. The fact that a simple change, in only one allele, can underlie such a distinctive phenotype as pesticide resistance is a remarkable lesson for molecular toxicology.

A recent study on killifish, *Fundulus heteroclitus*, along the East coast of the United States has revealed a much more complicated pattern of resistance. Populations of these fish live in estuaries, some with severely polluted sediments, containing high concentrations of polychlorinated biphenyls (PCBs) and polycyclic aromatic hydrocarbons (PAHs). Killifish from the polluted environments are much more resistant to toxicity from the model compounds PCB126 and benzo(a)pyrene. This resistance is related to mutations in the gene encoding aryl hydrocarbon receptor (AHR), the protein that binds PAHs and certain PCB metabolites and activates CYP expression. Also mutations in a protein called aryl-hydrocarbon receptor-interacting protein (AIP), a protein that combines with AHR to ensure binding of the ligand, contribute to down-regulation of the CY1A1 pathway. The net result is that killifish CYP1A1 shows only moderate induction by PCBs and PAHs and the damaging effects of reactive metabolites are avoided. However, since direct knockdown of CYP1A1 does not provide resistance it is still unclear whether the beneficial effects of the mutations in AHR actually act through an effect on CYP1A1.

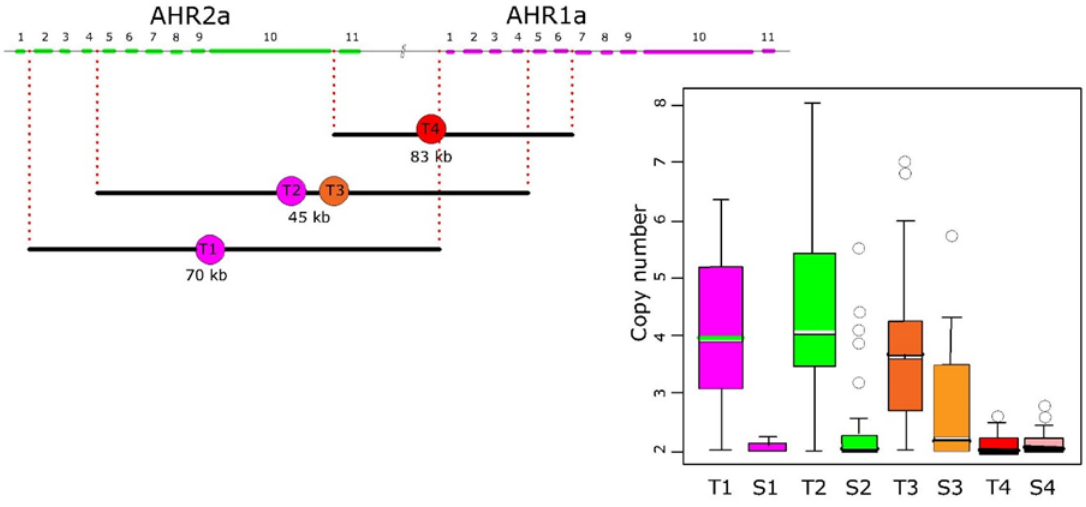


Figure 1. Showing the genetic variation among sensitive (S1 to S4) and tolerant (T1 to T4) populations of killifish, Fundulus heteroclitus along the East coast of the United States. Sensitivity and tolerance is towards sediments with high loads of PCBs and/or PAHs. The genome of Fundulus encodes four AHR (aryl hydrocarbon receptor) paralogs of which two are positioned in tandem, AHR2a and AHR1a, which carry long deletions (three different ones), indicated by black bars in the left figure. In addition, the populations have variable number of duplications of the CYP1A1 genes (right figure), not present to the same degree in the sensitive populations. Knock-out of AHR2a is protective of PCB and PAH toxicity, while duplication of CYP1A1 ensures a basal gene dose even when induction is less strong. Redrawn from Reid et al. (2016) by Wilma Ijzerman.

Interestingly, the various killifish populations show at least three different deletions in the AHR genes (Figure 1). In addition, the tolerant populations show various degrees of CYP1A1 duplication; in one population even eight paralogs are present. This can be interpreted as compensatory adaptations ensuring a basal constitutive level of CYP1A1 protein to conduct routine metabolic

activities. The killifish example shows a wonderful case of interplay between genetic tinkering, and strong selection emanating from a polluted environment.

Conclusion

In this module we have focused on genetic variation in the phase I enzyme, cytochrome P450. A similar complexity lies behind the phase II enzymes and the various xenobiotic-induced transporters (phase III). Still the P450 examples suffice to demonstrate that the machinery of xenobiotic metabolism shows a very large degree of genetic variation, as well as species differences due to duplications, deletions, gene conversion and lineage-specific selection. The variation resides both in copy number variation, alteration of coding sequences and in promoter or enhancer sequences affecting the expression of the enzymes. Such genetic variation is the template for evolution. In polluted environments enhanced expression is sometimes selected for (to neutralize toxic compounds), but sometimes also attenuated expression is selected (to avoid production of toxic intermediates). In the human genome, many of the polymorphisms have a medical significance, determining a personal profile of drug metabolism and tendencies to develop cancer.

References

Bell, G. (2012). Evolutionary rescue and the limits of adaptation. *Philosophical* Transactions of the Royal Society B 368, 2012.0080.

Daborn, P.J., Yen, J.L., Bogwitz, M.R., Le Goff, G., Feil, E., Jeffers, S., Tijet, N., Perry, T., Heckel, D., Batterham, P., Feyereisen, R., Wilson, T.G., Ffrench-Constant, R.H. (2002). A single P450 allele associated with insecticide resistance in *Drosophila*. Science 297, 2253-2256.

Feyereisen, R. (1999). Insect P450 enzymes. Annual Review of Entomology 44, 507-533.

Goldstone, H.M.H., Stegeman, J.J. (2006). A revised evolutionary history of the CYP1A subfamily: gene duplication, gene conversion and positive selection. Journal of Molecular Evolution 62, 708-717.

Kubota, A., Iwata, H., Goldstone, H.M.H., Kim, E.-Y., Stegeman, J.J., Tanabe, S. (2006). Cytochrome P450 1A1 and 1A5 in common cormorant (*Phalacrocorax carbo*): evolutionary relationships and functional implications associated with dioxin and related compounds. Toxicological Sciences 92, 394-408.

Reid, N.M., Proestou, D.A., Clark, B.W., Warren, W.C., Colbourne, J.K., Shaw, J.R., Hahn, M., Nacci, D., Oleksiak, M.F., Crawford, D.L., Whitehead, A. (2016). The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish Science 354, 1305-1308.

Preissner, S.C., Hoffmann, M.F., Preissner, R., Dunkel, R., Gewiess, A., Preissner, S. (2013). Polymorphic cytochrome P450 enyzmes (CYPs) and their role in personalized therapy. PLoS ONE 8, e82562.

Roszak, A., Lianeri, M., Sowinska, A., Jagodzinski, P.P. (2014). *CYP1A1* Ile462Val polymorphism as a risk factor in cervical cancer development in the Polish populations. Molecular Diagnosis and Therapy 18, 445-450.

Taylor, M., Feyereisen, R. (1996). Molecular biology and evolution of resistance to toxicants. Molecular Biology and Evolution 13, 719-734.

Walker, C.H., Ronis, M.J. (1989). The monooxygenases of birds, reptiles and amphibians. Xenobiotica 19, 1111-1121.

Zhou, S.-F., Liu J.-P., Chowbay, B. (2009). Polymorphism of human cytochrome P450 enzymes and its clinical impact. Drug Metabolism Reviews 41, 89-295.

4.2.14. Question 1

Comment upon: "In the human cytochrome P450 gene 1A1 133 single nucleotide polymorphisms have been described, of which 23 are non-synonymous".

- What is a "single nucleotide polymorphism" (SNP)?
- What is the difference between a synonymous SNP and a non-synonymous SNP?
- What could be the differences of nucleotide substitutions (1) in the promoter of CYP1A1, (2) in one of the introns of the CYP1A1 gene, (3) in one of the exons of CYP1A1?

- Does a single amino acid change in the cytochrome P450 protein have any medical relevance?

4.2.14. Question 2

Some populations of killifish along the Atlantic coast of the U.S. show very high resistance against organic contaminants such as dioxin-like PCBs. Genetic research has shown that these resistant populations have a deletion in the gene encoding aryl hydrocarbon receptor (Ahr). Explain why such a mutation can cause resistance to dioxin-like PCBs.

4.2.14. Question 3

Comment upon the concept of "Evolutionary rescue from pollution"- the evolution of tolerance as a way to diminish toxic effects of pollution.

This page titled 2.3.4.2: Toxicodynamics and Molecular Interactions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.2: Toxicodynamics and Molecular Interactions** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.4.3: Toxicity Testing

4.3. Toxicity testing

Author: Kees van Gestel

Reviewer: Michiel Kraak

Learning objectives:

You should be able to

- Mention the two general types of endpoints in toxicity tests
- Mention the main groups of test organisms used in environmental toxicology
- Mention different criteria determining the validity of toxicity tests
- Explain why toxicity testing may need a negative and a positive control

Keywords: single-species toxicity tests, test species selection, concentration-response relationships, endpoints, bioaccumulation testing, epidemiology, standardization, quality control, transcriptomics, metabolomics,

Introduction

Laboratory toxicity tests may provide insight into the potential of chemicals to bioaccumulate in organisms and into their hazard, the latter usually being expressed as toxicity values derived from concentration-response relationships. Section 4.3.1 on Bioaccumulation testing describes how to perform tests to assess the bioaccumulation potential of chemicals in aquatic and terrestrial organisms, and under static and dynamic exposure conditions. Basic to toxicity testing is the establishment of a concentration-response relationship, which relates the endpoint measured in the test organisms to exposure concentrations. Section 4.3.2 on Concentration-response relationships elaborates on the calculation of the relevant toxicity parameters like the median lethal concentration (LC_{50}) and the medium effective concentration (EC_{50}) from such toxicity tests. It also discusses the pros and cons of different methods for analyzing data from toxicity tests.

Several issues have to be addressed when designing toxicity tests that should enable assessing the environmental or human health hazard of chemicals. This concerns among others the selection of test organisms (see section 4.3.4 on the Selection of test organisms for ecotoxicity testing), exposure media, test conditions, test duration and endpoints, but also requires clear criteria for checking the quality of toxicity tests performed (see below). Different whole organism endpoints that are commonly used in standard toxicity tests, like survival, growth, reproduction or avoidance behavior, are discussed in section 4.3.3 on Endpoints. The sections 4.3.4 to 4.3.7 are focusing on the selection and performance of tests with organisms representative of aquatic and terrestrial ecosystems. This includes microorganisms (section 4.3.6), plants (section 4.3.5), invertebrates (section 4.3.4) and vertebrate test organisms (e.g. fish: section 4.3.4 on ecotoxicity tests, and birds: section 4.3.7). Testing of vertebrates, including fish (section 4.3.4) and birds (section 4.3.7), is subject to strict regulations, aimed at reducing the use of test animals. Data on the potential hazard of chemicals to human health therefore preferably have to be obtained in other ways, like by using *in vitro* test methods (section 4.3.8), by using data from post-registration monitoring of exposed humans (section 4.3.9 on Human toxicity testing), or from epidemiological analysis on exposed humans (section 4.3.10).

Inclusion of novel endpoints in toxicity testing

Traditionally, toxicity tests focus on whole organism endpoints, with survival, growth and reproduction being the most measured parameters (section 4.3.3). In case of vertebrate toxicity testing, also other endpoints may be used addressing effects at the level of organs or tissues (section 4.3.9 on human toxicity testing). Behavioural (e.g. avoidance behavior) and biochemical endpoints, like enzyme activity, are also regularly included in toxicity testing with vertebrates and invertebrates (sections 4.3.3, 4.3.4, 4.3.7, 4.3.9).

With the rise of molecular biology, novel techniques have become available that may provide additional information on the effects of chemicals. Molecular tools may, for instance, be applied in molecular epidemiology (section 4.3.11) to find causal relationships between health effects and the exposure to chemicals. Toxicity testing may also use gene expression responses (transcriptomics; section 4.3.12) or changes in metabolism (metabolomics; section 4.3.13) in relation to chemical exposures to help unraveling the mechanism(s) of action of chemicals. A major challenge still is to explain whole organism effects from such molecular responses.

Standardization of tests

The standardization of tests is organized by international bodies like the Organization for Economic Co-operation and Development (OECD), the International Standardization Organization (ISO), and ASTM International (formerly known as the American Society for Testing and Materials). Standardization aims at reducing variation in test outcomes by carefully describing the methods for culturing and handling the test organisms, the procedures for performing the test, the properties and composition of test media, the exposure conditions and the analysis of the data. Standardized test guidelines are usually based on extensive testing of a method by different laboratories in a so-called round-robin test.

Regulatory bodies generally require that toxicity tests supporting the registration of new chemicals are performed according to internationally standardized test guidelines. In Europe, for instance, all toxicity tests submitted within the framework of REACH have to be performed according to the OECD guidelines for the testing of chemicals (see section on Regulation of chemicals).

Quality control of toxicity tests

Since toxicity tests are performed with living organisms, this inevitably leads to (biological) variation in outcomes. Coping with this variation requires the use of sufficient replication, careful test designs and good choice of endpoints (section 4.3.3) to enable proper estimates of relevant toxicity data.

In order to control the quality of the outcome of toxicity tests, several criteria have been developed, which mainly apply to the performance of the test organisms in the non-exposed controls. These criteria may e.g. require a minimum % survival of control organisms, a minimum growth rate or number of offspring being produced by the controls and limited variation (e.g. <30%) of the replicate control growth or reproduction data (sections 4.3.4, 4.3.5, 4.3.6, 4.3.7). When tests do not meet these criteria, the outcome is prone to doubts, as for instance a poor control survival will make it hard to draw sound conclusions on the effect of the test chemical on this endpoint. As a consequence, tests that do not meet these validity criteria may not be accepted by other scientists and by regulatory authorities.

In case the test chemical is added to the test medium using a solvent, toxicity tests should also include a solvent control, in addition to a regular non-exposed control (see section 4.3.4 on the selection of test organisms for ecotoxicity testing). In case the response in the solvent control differs significantly from that in the negative control, the solvent control will be used as the control for analyzing the effects of the test chemical. The negative control will then only be used to check if the validity criteria have been met and to monitor the condition of the test organisms. In case the responses in the negative control and the solvent control do not differ significantly, both controls can be pooled for the data analysis.

Most test guidelines also require frequent testing of a *positive control*, a chemical with known toxicity, to check if the long-term culturing of the test organisms does not lead to changes in their sensitivity.

4.3. Question 1

What are the main endpoints in toxicity testing?

4.3. Question 2

Which are the main groups of organisms used in toxicity testing?

4.3. Question 3

Why is standardization of methods for toxicity testing required?

4.3. Question 4

Which elements are included in the quality control of toxicity tests?

4.3.1. Bioaccumulation testing

Author: Kees van Gestel

Reviewers: Joop Hermens, Michiel Kraak, Susana Loureiro

Learning objectives:

You should be able to

- describe methods for determining the bioaccumulation of chemicals in terrestrial and aquatic organisms
- describe a test design suitable for assessing the bioaccumulation kinetics of chemicals in organisms
- mention the pros and cons of static and dynamic bioaccumulation tests

Keywords: bioconcentration, bioaccumulation, uptake and elimination kinetics, test methods, soil, water

Bioaccumulation is defined as the uptake of chemicals in organisms from the environment. The degree of bioaccumulation is usually indicated by the bioconcentration factor (BCF) in case the exposure is via water, or the biota-to-soil/sediment accumulation factor (BSAF) for exposure in soil or sediment (see section on Bioaccumulation).

Because of the potential risk for food-chain transfer, experimental determination of the bioaccumulation potential of chemicals is usually required in case of a high lipophilicity (log K_{ow} > 3), unless the chemical has a very low persistency. For very persistent chemicals, experimental determination of bioaccumulation potential may already be triggered at log K_{ow} > 2. The experimental determination of BCF and BSAF values makes use of static or dynamic exposure systems.

In *static tests*, the medium is dosed once with the test chemical, and organisms are exposed for a certain period of time after which both the organisms and the test medium are analyzed for the test chemical. The BCF or BSAF are calculated from the measured concentrations. There are a few concerns with this way of bioaccumulation testing.

First, exposure concentrations may decrease during the test, e.g. due to (bio)degradation, volatilization, sorption to the walls of the test container, or uptake of the test compound by the test organisms. As a consequence, the concentration in the test medium measured at the start of the test may not be indicative of the actual exposure during the test. To take this into account, exposure concentrations can be measured at the start and the end of the test and also at some intermediate time points. Body concentrations in the test organisms may then be related to *time-weighted-average (TWA) exposure concentrations*. Alternatively, to overcome the problem of decreasing concentrations in aquatic test systems, continuous flow systems or passive dosing techniques can be applied. Such methods, however, are not applicable to soil or sediment tests, where repeated transfer of organisms to freshly spiked medium is the only way to guarantee more or less constant exposure concentrations in case of rapidly degrading compounds. To avoid that the uptake of the test chemical in test organisms leads to decreasing exposure concentrations, the amount of biomass per volume or mass of test medium should be sufficiently low.

Second, it is uncertain whether at the end of the exposure period steady state or equilibrium is reached. If this is not the case, the resulting BSAF or BCF values may underestimate the bioaccumulation potential of the chemical. To tackle this problem, a dynamic test may be run to assess the uptake and elimination rate constants to derive a BSAF or BCF values using uptake and elimination rate constants (see below).

Such uncertainties also apply to BCF and BSAF values obtained by analyzing organisms collected from the field and comparing body concentrations with exposure levels in the environment. Using data from field-exposed organisms on one hand have large uncertainty as it remains unclear whether equilibrium was reached, on the other hand they to do reflect exposure over time under fluctuating but realistic exposure conditions.

Dynamic tests, also indicated as uptake/elimination or toxicokinetic tests, may overcome some, but not all, of the disadvantages of static tests. In dynamic tests, organisms are exposed for a certain period of time in spiked medium to assess the uptake of the chemical, after which they are transferred to clean medium for determining the elimination of the chemical. During both the uptake and the elimination phase, at different points in time, organisms are sampled and analyzed for the test chemical. The medium is also sampled frequently to check for a possible decline of the exposure concentration during the uptake phase. Also in dynamic tests, keeping exposure concentrations constant as much as possible is a major challenge, requiring frequent renewal (see above).

Toxicokinetic tests should also include controls, consisting of test organisms incubated in the clean medium and transferred to clean medium at the same time the organisms from the treated medium are transferred. Such controls may help identifying possible irregularities in the test, such as poor health of the test organisms or unexpected (cross)contamination occurring during the test.

The concentrations of the chemical measured in the test organisms are plotted against the exposure time, and a first-order one-compartment model is fitted to the data to estimate the uptake and elimination rate constants. The (dynamic) BSAF or BCF value is then determined as the ratio of the uptake and elimination rate constants (see section on Bioconcentration and kinetic models).

In a toxicokinetics test, usually replicate samples are taken at each point in time, both during the uptake and the elimination phase. The frequency of sampling may be higher at the beginning than at the end of both phases: a typical sampling scheme is shown in Figure 1. Since the analysis of toxicokinetics data using the one-compartment model is regression based, it is generally preferred to have more points in time rather than having many replicates per sampling time. From that perspective, often no more than 3-4 replicates are used per sampling time, and 5-6 sampling times for the uptake and elimination phases each.

UPTAKE PHASE

Day	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Animals	3	3	3		3			3			3				3			3				3
Medium	X							X							X							X

ELIMINATION PHASE

Day	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Animals	3	3	3		3			3			3				3			3				3
Medium	X																					X

Figure 1. Sampling scheme of a toxicokinetics test for assessing the uptake and elimination kinetics of chemicals in earthworms. During the 21-day uptake phase, the earthworms are individually exposed to a test chemical in soil, and at regular intervals three earthworms are sampled. After 21 days, the remaining earthworms are transferred to clean soil for the 21-day elimination period, in which again three replicate earthworms are sampled at regular points in time for measuring the body concentrations of the chemical. Also the soil is analyzed at different points in time (marked with X in the Medium row). Drawn by the author.

Preferably, replicates are independent, so destructively sampled at a specific sampling point. Especially in aquatic ecotoxicology, mass exposures are sometimes used, having all test organisms in one or few replicate test containers. In this case, at each sampling time some replicate organisms are taken from the test container(s), and at the end of the uptake phase all organisms are transferred to (a) container(s) with clean medium.

Figure 2 shows the result of a test on the uptake and elimination kinetics of molybdenum in the earthworm *Eisenia andrei*. From the ratio of the uptake rate constant (k_1) and elimination rate constant (k_2) a BSAF of approx. 1.0 could be calculated, suggesting a low bioaccumulation potential of Mo in earthworms in the soil tested.

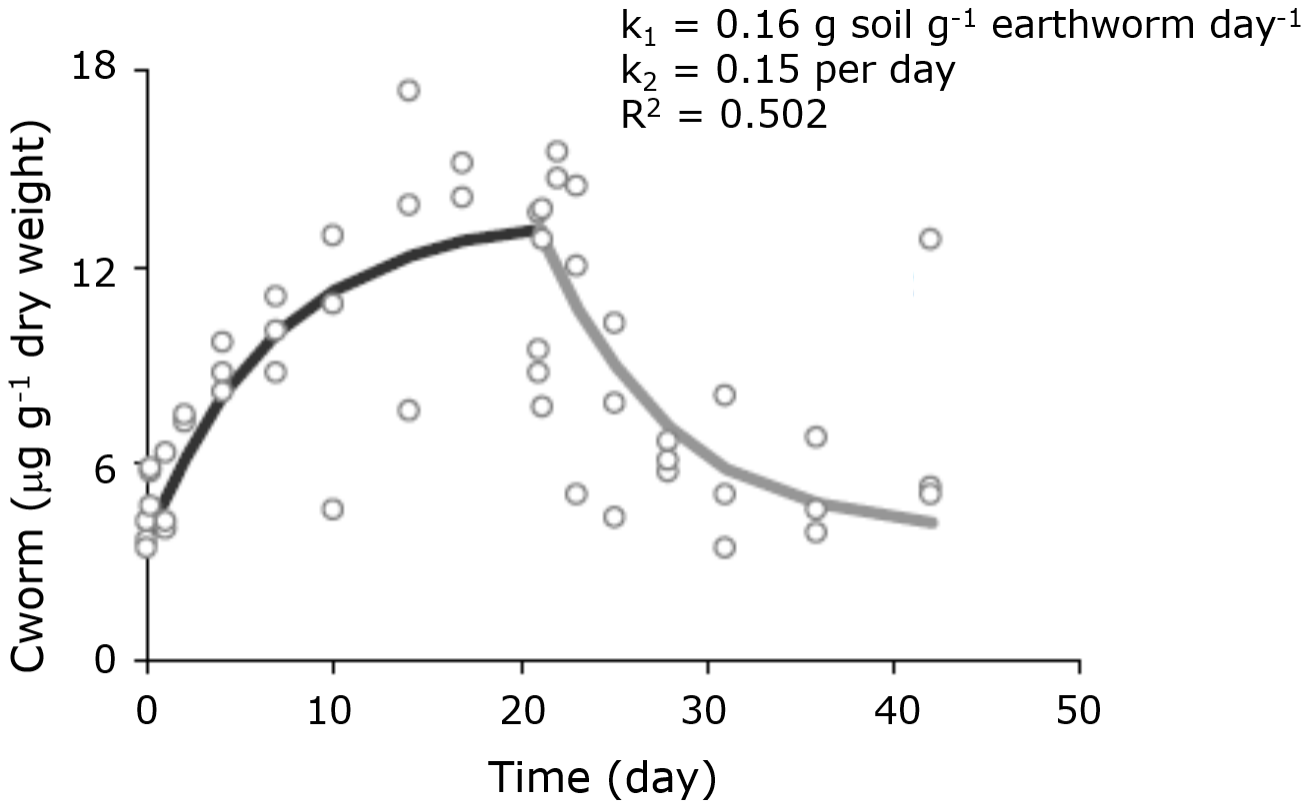


Figure 2. Uptake and elimination kinetics of molybdenum in Eisenia andrei exposed in an artificial soil spiked with a nominal Mo concentration of 10 µg g^{-1} dry soil. Dots represent measured internal Mo concentrations. Curves were estimated by a one-compartment model (see section on Bioconcentration and kinetic models). Parameters: k_1 = uptake rate constant [$g_{soil}/g_{worm}/d$], k_2 = elimination rate constant [d^{-1}]. Adapted from Diez-Ortiz et al. (2010).

Another way of assessing the bioaccumulation potential of chemicals in organisms includes the use of radiolabeled chemicals, which may facilitate easy detection of the test chemical. The use of radiolabeled chemicals may however, overestimate bioaccumulation potential when no distinction is made between the parent compound and potential metabolites. In case of metals, stable isotopes may also offer an opportunity to assess bioaccumulation potential. Such an approach was also applied to distinguish the role of dissolved (ionic) Zn in the bioaccumulation of Zn in earthworms from ZnO nanoparticles. Earthworms were exposed to soils spiked with mixtures of $^{64}ZnCl_2$ and ^{68}ZnO nanoparticles. The results showed that dissolution of the nanoparticles was fast and that the earthworms mainly accumulated Zn present in ionic form in the soil solution (Laycock et al., 2017).

Standard test guidelines for assessing the bioaccumulation (kinetics) of chemicals have been published by the Organization for Economic Cooperation and Development (OECD) for sediment-dwelling oligochaetes (OECD, 2008), for earthworms/enchytraeids in soil (OECD, 2010) and for fish (OECD, 2012).

References

Diez-Ortiz, M., Giska, I., Groot, M., Borgman, E.M., Van Gestel, C.A.M. (2010). Influence of soil properties on molybdenum uptake and elimination kinetics in the earthworm *Eisenia andrei*. Chemosphere 80, 1036-1043.

Laycock, A., Romero-Freire, A., Najorka, J., Svendsen, C., Van Gestel, C.A.M., Rehkämper, M. (2017). Novel multi-isotope tracer approach to test ZnO nanoparticle and soluble Zn bioavailability in joint soil exposures. Environmental Science and Technology 51, 12756−12763.

OECD (2008). Guidelines for the testing of chemicals No. 315: Bioaccumulation in Sediment-dwelling Benthic Oligochaetes. Organization for Economic Cooperation and Development, Paris.

OECD (2010). Guidelines for the testing of chemicals No. 317: Bioaccumulation in Terrestrial Oligochaetes. Organization for Economic Cooperation and Development, Paris.

OECD (2012). Guidelines for the testing of chemicals No. 305: Bioaccumulation in Fish: Aqueous and Dietary Exposure. Organization for Economic Cooperation and Development, Paris.

4.3.1. Question 1

Why may BSAF and BCF values obtained from static tests not reflect the real bioaccumulation potential of chemicals, even if it was possible to keep exposure concentrations constant?

4.3.1. Question 2

Describe the experimental design of a test for assessing the uptake and elimination kinetics of chemicals in test organisms, in soil or water.

4.3.1. Question 3

a. What experimental problem may be encountered when determining the bioaccumulation of chemicals in terrestrial or aquatic organisms?

b. And how may this problems be overcome in case of aquatic organisms?

c. Is such a solution also possible for terrestrial organisms?

4.3.2. Concentration-response relationships

Author: Kees van Gestel

Reviewers: Michiel Kraak, Thomas Backhaus

Learning goals:

You should be able to

- understand the concept of the concentration-response relationship
- define measures of toxicity
- distinguish quantal and continuous data
- mention the reasons for preferring ECx values above NOEC values

Keywords: concentration-related effects, measure of lethal effect, measure of sublethal effect, regression-based analysis

Key paradigm in human and environmental toxicology is that the dose determines the effect. This paradigm goes back to Paracelsus, stating that any chemical is toxic, but that the dose determines the severity of the effect. In practice, this paradigm is used to quantify the toxicity of chemicals. For that purpose, toxicity tests are performed in which organisms (microbes, plants, invertebrates, vertebrates) or cells are exposed to a range of concentrations of a chemical. Such tests also include incubations in non-treated control medium. The response of the test organisms is determined by monitoring selected endpoints, like survival, growth, reproduction or other parameters (see section on Endpoints). Endpoints can increase (e.g. mortality) or decrease with increasing exposure concentration (e.g. survival, reproduction, growth). The response of the endpoints is plotted against the exposure concentration, and so-called concentration-response curves (Figure 1) are fitted, from which measures of the toxicity of the chemical can be calculated.

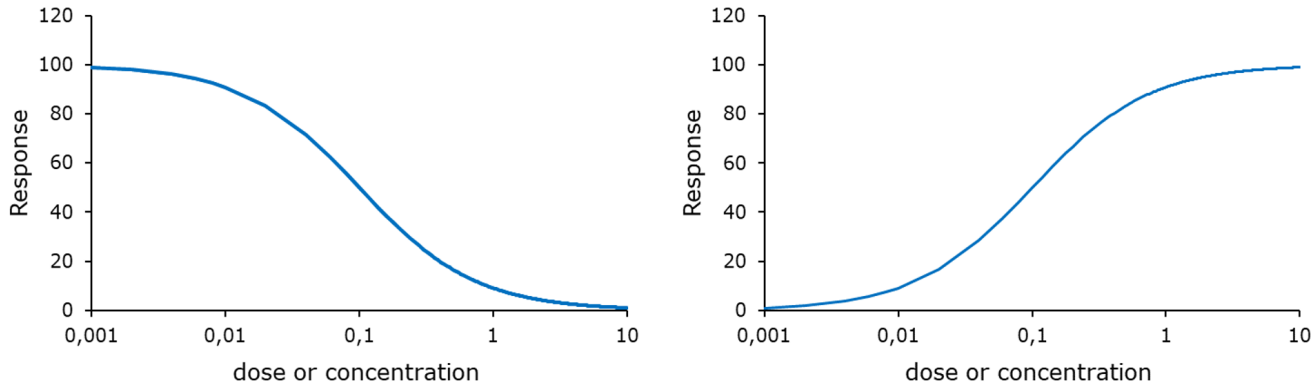


Figure 1: Concentration-response relationships. Left: response of the endpoint (e.g., survival, reproduction, growth) decreases with increasing concentration. Right: response of the endpoint (e.g., mortality, induction of enzyme activity) increases with increasing exposure concentration.

The unit of exposure, the concentration or dose, may be expressed differently depending on the exposed subject. Dose is expressed as mg/kg body weight in human toxicology and following single (oral or dermal) exposure events in mammals or birds. For other orally or dermally exposed (invertebrate) organisms, like honey bees, the dose may be expressed per animal, e.g. µg/bee. Environmental exposures generally express exposure as the concentration in mg/kg food, mg/kg soil, mg/l surface, drinking or ground water, or mg/m^3 air.

Ultimately, it is the concentration (number of molecules of the chemical) at the target site that determines the effect. Consequently, expressing exposure concentrations on a molar basis (mol/L, mol/kg) is preferred, but less frequently applied.

At low concentrations or doses, the endpoint measured is not affected by exposure. At increasing concentration, the endpoint shows a concentration-related decrease or increase. From this decrease or increase, different measures of toxicity can be calculated:

ECx/EDx: the "effective concentration" resp. "effective dose"; "x" denotes the percentage effect relative to an untreated control. This should always be followed by giving the selected endpoint.

LCx/LDx: same, but specified for a specific endpoint: lethality.

EC50/ED50: the median effect concentration or dose, with "x" set to 50%. This is the most common estimate used in environmental toxicology. This should always be followed by giving the selected endpoint.

LC50/LD50: same, but specified for a specific endpoint: lethality.

The terms LCx and LDx refer to the fraction of animals responding (dying), while the ECx and EDx indicate the degree of reduction of the measured parameter. The ECx/EDx describe the overall average performance of the test organisms in terms of the parameter measured (e.g., growth, reproduction). The meaning of an LCx/LDx seems obvious: it refers to lethality of the test chemical. The use of ECx/EDx, however, always requires explicit mentioning of the endpoint it concerns.

Concentration-response models usually distinguish *quantal* and *continuous data*. Quantal data refer to constrained ("yes/no") responses and include, for instance, survival data, but may also be applicable to avoidance responses. Continuous data refer to parameters like growth, reproduction (number of juveniles or eggs produced) or biochemical and physiological measurements. A crucial difference between quantal and continuous responses is that quantal responses are population-level responses, while continuous responses can also be observed on the level of individuals. An organism cannot be half-dead, but it can certainly grow at only half the control rate.

Concentration-response models are usually sigmoidal on a log-scale and are characterized by four parameters: *minimum*, *maximum*, *slope* and *position*. The *minimum response* is often set to the control level or to zero. The *maximum response* is often set to 100%, in relation to the control or the biologically plausible maximum (e.g. 100% survival). The *slope* identifies the steepness of the curve, and determines the distance between the EC50 and EC10. The *position* parameter indicates where on the x-axis the curve is placed. The position may equal the EC50 and in that case it is named the *turning point*. But this in fact holds only for a small fraction of models and not for models that are not symmetrical to the EC50.

In environmental toxicology, the parameter values are usually presented with *95% confidence intervals* indicating the margins of uncertainty. Statistical software packages are used to calculate these corresponding 95% confidence intervals.

Regression-based test designs require several test concentrations, and the results are dependent on the used statistical model, especially in the low-effect region. Sometimes it is simply impossible to use a regression-based design because the endpoint does not cover a sufficiently high effect range (>50% effect is typically needed for an accurate fit).

In case of quantal responses, especially survival, the slope of the concentration-response curve is an indication of the *sensitivity distribution* of the individuals within the population of test organisms. For a very homogenous population of laboratory test animals having the same age and body size, a steeper concentration-response curve is expected than when using field-collected animals representing a wider range of ages and body sizes (Figure 2).

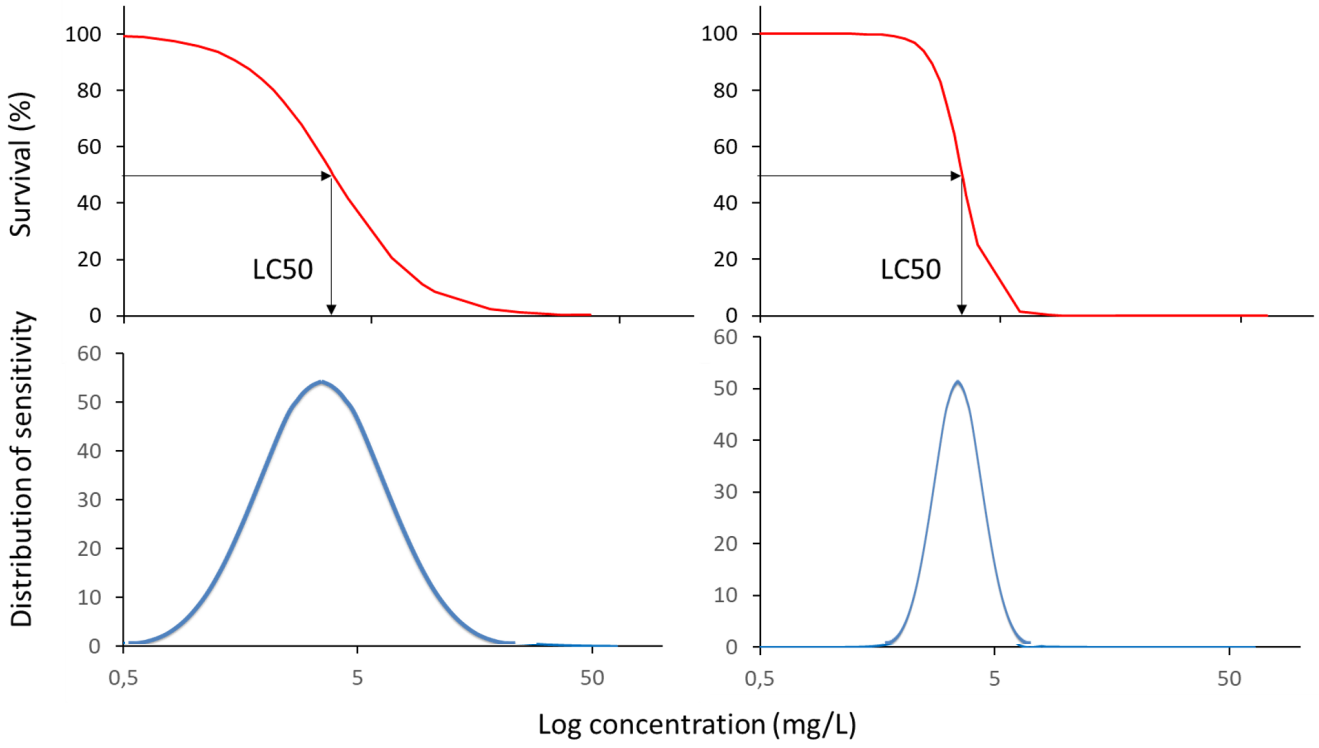


Figure 2: The steepness of the concentration-response curve for effects on survival (top) may provide insight into the sensitivity distribution of the individuals within the population of test animals (bottom). The steeper the curve the smaller the variation in sensitivity among the test organisms.

In addition to ECx values, toxicity tests may also be used to derive other measures of toxicity:

NOEC/NOEL: No-Observed Effect Concentration or Effect Level

LOEC/LOEL: Lowest Observed Effect Concentration or Effect Level

NOAEL: No-Observed Adverse Effect Level. Same as NOEL, but focusing on effects that are negative (adverse) compared to the control.

LOAEL: Lowest Observed Adverse Effect Level. Same as LOEL, but focusing on effects that are negative (adverse) compared to the control.

Where the ECx are derived by curve fitting, the NOEC and LOEC are derived by a statistical test comparing the response at each test concentration with that of the controls. The NOEC is defined as the highest test concentration where the response does not significantly differ from the control. The LOEC is the next higher concentration, so the lowest concentration tested at which the response significantly differs from the control. Figure 3 shows NOEC and LOEC values derived from a hypothetical test. Usually an Analysis of Variance (ANOVA) is used combined with a post-hoc test, e.g. Tukey, Bonferroni or Dunnett, to determine the NOEC and LOEC.

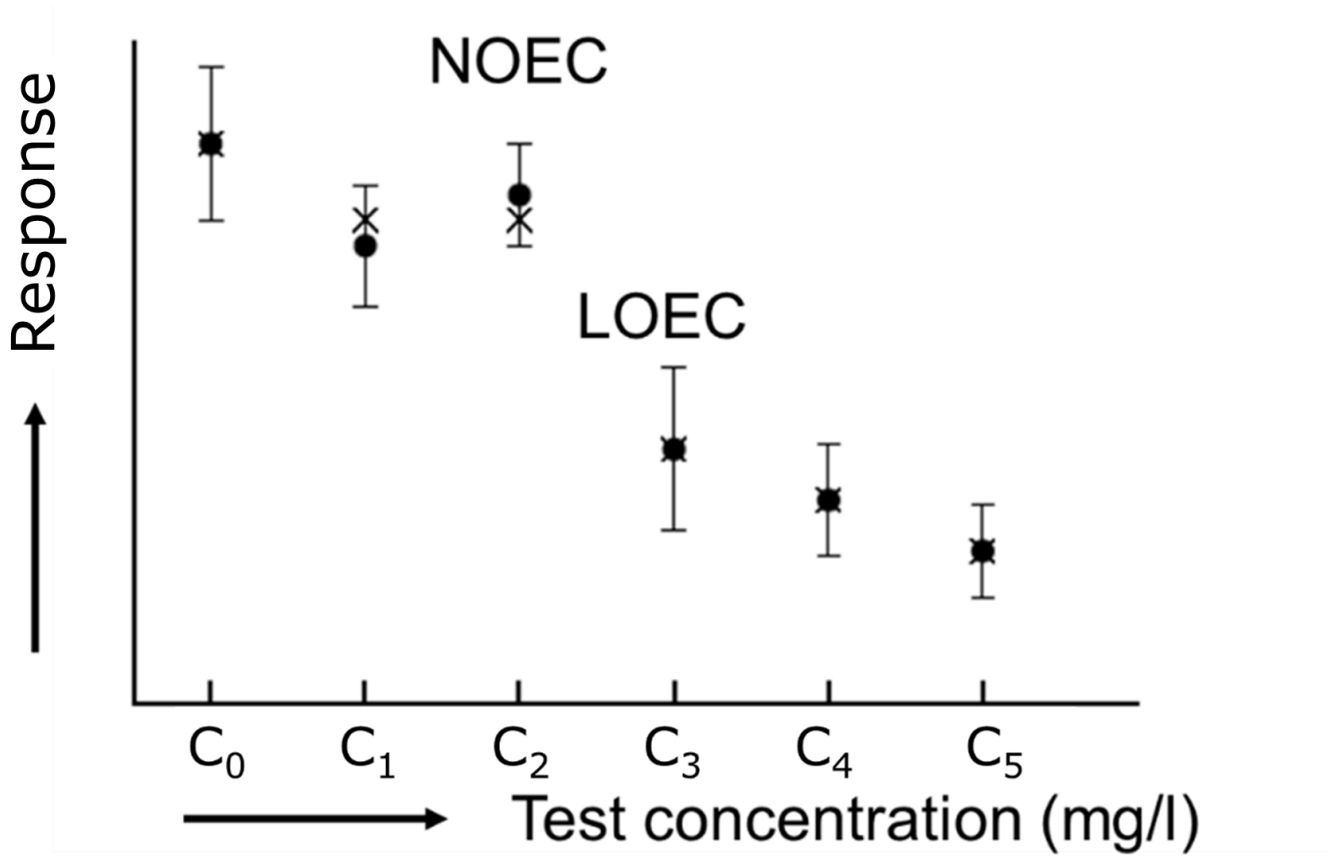


Figure 3: Derivation of NOEC and LOEC values from a toxicity test.

Most available toxicity data are NOECs, hence they are the most common values found in databases and therefore used for regulatory purposes. From a scientific point of view, however, there are quite some disadvantages related to the use of NOECs:

- Obtained by statistical test (hypothesis testing) (compared to regression analysis);
- Equal to one of the test concentrations, so not using all data from the toxicity test;
- Sensitive to the number of replicates used per exposure concentration and control;
- Sensitive to variation in response, so for differences between replicates;
- Depends on the statistical test chosen, and on the variance (σ);
- Does not have confidence intervals;
- Makes it hard to compare toxicity data between laboratories and between species.

The NOEC may, due to its sensitivity to variation and test design, sometimes be equal to or even higher than the EC50.

Because of the disadvantages of the NOEC, it is recommended to use measures of toxicity derived by fitting a concentration-response curve to the data obtained from a toxicity test. As an alternative to the NOEC, usually an EC10 or EC20 is used, which has the advantages that it is obtained using all data from the test and that it has a 95% confidence interval indicating its reliability. Having a 95% confidence interval also allows a statistical comparison of ECx values, which is not possible for NOEC values.

4.3.2. Question 1

Which four parameters describe a dose-response curve?

4.3.2. Question 2

What would be the preferred unit of measures of toxicity (e.g. EC20 or EC50) describing the effect of a chemical on the survival of soil invertebrates exposed in a standardized test soil?

4.3.2. Question 3

Why would you expect that using an age-synchronized laboratory population of test organisms results in a much steeper concentration-response curve for effects on survival of a chemical than a field-collected population of non-synchronized individuals?

4.3.2. Question 4

Why are EC10 values preferred over NOECs when using the outcomes of toxicity tests for the risk assessment of chemicals?

4.3.3. Endpoints

Author: Michiel Kraak

Reviewers: Kees van Gestel, Carlos Barata

Learning objectives:

You should be able to

- list the available whole organism endpoints in toxicity tests.
- motivate the importance of sublethal endpoints in acute and chronic toxicity tests.
- describe how sublethal endpoints in acute and chronic toxicity tests are measured.

Keywords: Mortality, survival, sublethal endpoints, growth, reproduction, behaviour, photosynthesis

Introduction

Most toxicity tests performed are short-term high-dose experiments, acute tests in which mortality is often the only endpoint. Mortality, however, is a crude parameter in response to relatively high and therefore often environmentally irrelevant toxicant concentrations. At much lower and therefore environmentally more relevant toxicant concentrations, organisms may suffer from a wide variety of sublethal effects. Hence, toxicity tests gain ecological realism if sublethal endpoints are addressed in addition to mortality.

Mortality

Mortality can be determined in both acute and chronic toxicity tests. In acute tests, mortality is often the only feasible endpoint, although some acute tests take long enough to also measure sublethal endpoints, especially growth. Generally though, this is restricted to chronic toxicity tests, in which a wide variety of sublethal endpoints can be assessed in addition to mortality (Table 1).

Mortality at the end of the exposure period is assessed by simply counting the number of surviving individuals, but it can also be expressed either as percentage of the initial number of individuals or as percentage of the corresponding control. The increasing mortality with increasing toxicant concentrations can be plotted in a dose-response relationship from which the LC50 can be derived (see section on Concentration-response relationship). If assessing mortality is non-destructive, for instance if this can be done by visual inspection, it can be scored at different time intervals during a toxicity test. Although repeated observations may take some effort, they generally do generate valuable insights in the course of the intoxication process over time.

Sublethal endpoints in acute toxicity tests

In acute toxicity tests it is difficult to assess other endpoints than mortality, since effects of toxicants on sublethal endpoints like growth and reproduction need much longer exposure times to become expressed (see section on Chronic toxicity). Incorporating sublethal endpoints in acute toxicity tests thus requires rapid responses to toxicant exposure. Photosynthesis of plants and behaviour of animals are elegant, sensitive and rapidly responding endpoints that can be incorporated into acute toxicity tests (Table 1).

Behavioural endpoints

Behaviour is an understudied but sensitive and ecologically relevant endpoint in ecotoxicity testing, since subtle changes in animal behaviour may affect trophic interactions and ecosystem functioning. Several studies reported effects on animal behaviour at concentrations orders of magnitudes lower than lethal concentrations. Van der Geest et al. (1999) showed that changes in ventilation behaviour of fifth instar larvae of the caddisfly *Hydropsyche angustipennis* occurred at approximately 150 times lower

Cu concentrations than mortality of first instar larvae. Avoidance behaviour of the amphipod *Corophium volutator* to contaminated sediments was 1,000 times more sensitive than survival (Hellou et al., 2008). Chevalier et al. (2015) tested the effect of twelve compounds covering different modes of action on the swimming behaviour of daphnids and observed that most compounds induced an early and significant swimming speed increase at concentrations near or below the 10% effective concentration (48-h EC_{10}) of the acute immobilization test. Barata et al. (2008) reported that the short term (24 h) *D. magna* feeding inhibition assay was on average 50 times more sensitive than acute standardized tests when assessing the toxicity of a mixture of 16 chemicals in different water types combinations. These and many other examples all show that organisms may exhibit altered behaviour at relatively low and therefore often environmentally relevant toxicant concentrations.

Behavioural responses to toxicant exposure can also be very fast, allowing organisms to avoid further exposure and subsequent bioaccumulation and toxicity. A wide array of such avoidance responses have been incorporated in ecotoxicity testing (Araújo et al., 2016), including the avoidance of contaminated soil by earthworms (*Eisenia fetida*) (Rastetter & Gerhardt; 2018), feeding inhibition of mussels (*Corbicula fluminea*) (Castro et al., 2018), aversive swimming response to silver nanoparticles by the unicellular green alga *Chlamydomonas reinhardtii* (Mitzel et al., 2017) and by daphnids to twelve compounds covering different modes of toxic action (Chevalier et al., 2015).

Photosynthesis

Photosynthesis is a sensitive and well-studied endpoint that can be applied to identify hazardous effects of herbicides on primary producers. In bioassays with plants or algae, photosynthesis is often quantified using pulse amplitude modulation (PAM) fluorometry, a rapid measurement technique suitable for quick screening purposes. Algal photosynthesis is preferably quantified in light adapted cells as effective photosystem II (PSII) efficiency (ΦPSII) (Ralph et al., 2007; Sjollema et al., 2014). This endpoint responds most sensitively to herbicide activity, as the most commonly applied herbicides either directly or indirectly affect PSII (see section on Herbicide toxicity).

Sublethal endpoints in chronic toxicity tests

Besides mortality, growth and reproduction are the most commonly assessed endpoints in ecotoxicity tests (Table 1). Growth can be measured in two ways, as an increase in length and as an increase in weight. Often only the length or weight at the end of the exposure period is determined. This, however, includes both the growth before and during exposure. It is therefore more distinctive to measure length or weight at the beginning as well as at the end of the exposure, and then subtract the individual or average initial length or weight from the final individual length or weight. Growth during the exposure period may subsequently be expressed as percentage of the initial lengths or weight. Ideally the initial length or weight is measured from the same individuals that will be exposed. When organisms are sacrificed to measure the initial length or weight, which is especially the case for dry weight, this is not feasible. In that case a subsample from the individuals is taken apart at the beginning of the test.

Reproduction is a sensitive and ecological relevant endpoint in chronic toxicity tests. It is an integrated parameter, incorporating many different aspects of the process, that can be assessed one by one. The first reproduction parameter is the day of first reproduction. This is an ecologically very relevant parameter, as delayed reproduction obviously has strong implications for population growth. The next reproduction parameter is the amount of offspring. In this case the number of eggs, seeds, neonates or juveniles can be counted. For organisms that produce egg ropes or egg masses, both the number of egg masses as well as the number of eggs per mass can be determined. Lastly the quality of the offspring can be quantified. This can be achieved by determining their physiological status (e.g. fat content), their size, survival and finally their chance or reaching adulthood.

Table 1. *Whole organism endpoints often used in toxicity tests. Quantal refers to a yes/no endpoint, while graded refers to a continuous endpoint (see section on Concentration-response relationship).*

Endpoint	Acute/Chronic	Quantal/Graded
mortality	both	quantal
behaviour	acute	graded
avoidance	acute	quantal
photosynthesis	acute	graded
growth (length and weight)	mostly chronic	graded
reproduction	chronic	graded

A wide variety of other, less commonly applied sublethal whole organism endpoints can be assessed upon chronic exposure. The possibilities are endless, with some specific endpoints being designed for the effect of a single compound only, or species specific endpoints, sometimes described for only one organism. Sub-organismal endpoints are described in a separate chapter (see section on Molecular endpoints in toxicity tests).

References

Araujo, C.V.M., Moreira-Santos, M., Ribeiro, R. (2016). Active and passive spatial avoidance by aquatic organisms from environmental stressors: A complementary perspective and a critical review. Environment International 92-93, 405-415.

Barata, C., Alanon, P., Gutierrez-Alonso, S., Riva, M.C., Fernandez, C., Tarazona, J.V. (2008). A *Daphnia magna* feeding bioassay as a cost effective and ecological relevant sublethal toxicity test for environmental risk assessment of toxic effluents. Science of the Total Environment 405(1-3), 78-86.

Castro, B.B., Silva, C., Macario, I.P.E., Oliveira, B., Concalves, F., Pereira, J.L. (2018). Feeding inhibition in *Corbicula fluminea* (OF Muller, 1774) as an effect criterion to pollutant exposure: Perspectives for ecotoxicity screening and refinement of chemical control. Aquatic Toxicology 196, 25-34.

Chevalier, J., Harscoët, E., Keller, M., Pandard, P., Cachot, J., Grote, M. (2015). Exploration of *Daphnia* behavioral effect profiles induced by a broad range of toxicants with different modes of action. Environmental Toxicology and Chemistry 34, 1760-1769.

Hellou J., Cheeseman, K., Desnoyers, E., Johnston, D., Jouvenelle, M.L., Leonard, J., Robertson, S., Walker, P. (2008). A non-lethal chemically based approach to investigate the quality of harbor sediments. Science of the Total Environment 389, 178-187.

Mitzel, M.R., Lin, N., Whalen, J.K., Tufenkji, N. (2017). *Chlamydomonas reinhardtii* displays aversive swimming response to silver nanoparticles Environmental Science: Nano 4, 1328-1338.

Ralph, P.J., Smith, R.A., Macinnis-Ng, C.M.O., Seery, C.R. (2007). Use of fluorescence-based ecotoxicological bioassays in monitoring toxicants and pollution in aquatic systems: Review. Toxicological and Environmental Chemistry 89, 589-607.

Rastetter, N., Gerhardt, A. (2018). Continuous monitoring of avoidance behaviour with the earthworm *Eisenia fetida*. Journal of Soils and Sediments 18, 957-967.

Sjollema, S.B., Van Beusekom, S.A.M., Van der Geest, H.G., Booij, P., De Zwart, D., Vethaak, A.D., Admiraal, W. (2014). Laboratory algal bioassays using PAM fluorometry: Effects of test conditions on the determination of herbicide and field sample toxicity. Environmental Toxicology and Chemistry 33, 1017-1022.

Van der Geest, H.G., Greve, G.D., De Haas, E.M., Scheper, B.B., Kraak, M.H.S., Stuijfzand, S.C., Augustijn, C.H., Admiraal, W. (1999). Survival and behavioural responses of larvae of the caddisfly *Hydropsyche angustipennis* to copper and diazinon. Environmental Toxicology and Chemistry 18, 1965-1971.

4.3.3. Question 1

What is the importance of incorporating sublethal endpoints in acute and chronic toxicity tests?

4.3.3. Question 2

Name one animal and one plant specific sublethal endpoint that can be incorporated in acute toxicity test.

4.3.3. Question 3

Name the two most commonly assessed endpoints in chronic toxicity tests.

4.3.4. Selection of test organisms - Eco animals

Author: Michiel Kraak

Reviewers: Kees van Gestel, Jörg Römbke

Learning objectives:

You should be able to

- name the requirements for suitable laboratory ecotoxicity test organisms.
- list the most commonly used standard test organisms per environmental compartment.
- argue the need for more than one test species and the need for non-standard test organisms.

Key words: Test organism, standardized laboratory ecotoxicity tests, environmental compartment, habitat, different trophic levels

Introduction

Standardized laboratory ecotoxicity tests require constant test conditions, standardized endpoints (see section on Endpoints) and good performance in control treatments. Actually, in reliable, reproducible and easy to perform toxicity tests, the test compound should be the only variable. This sets high demands on the choice of the test organisms.

For a proper risk assessment, it is crucial that test species are representative of the community or ecosystem to be protected. Criteria for selection of organisms to be used in toxicity tests have been summarized by Van Gestel et al. (1997). They include: 1. Practical arguments, including feasibility, cost-effectiveness and rapidity of the test, 2. Acceptability and standardisation of the tests, including the generation of reproducible results, and 3. Ecological significance, including sensitivity, biological validity etc. The most practical requirement is that the test organism should be easy to culture and maintain, but equally important is that the test species should be sensitive towards different stressors. These two main requirements are, however, frequently conflicting. Species that are easy to culture are often less sensitive, simply because they are mostly generalists, while sensitive species are often specialists, making it much harder to culture them. For scientific and societal support of the choice of the test organisms, preferably they should be both ecologically and economically relevant or serve as flagship species, but again, these are opposite requirements. Economically relevant species, like crops and cattle, hardly play any role in natural ecosystems, while ecologically highly relevant species have no obvious economic value. This is reflected by the research efforts on these species, since much more is known about economically relevant species than about ecologically relevant species.

There is no species that is most sensitive to all pollutants. Which species is most sensitive depends on the mode of action and possibly also other properties of the chemical, the exposure route, its availability and the properties of the organism (e.g., presence of specific targets, physiology, etc.). It is therefore important to always test a number of species, with different life traits, functions, and positions in the food web. According to Van Gestel et al. (1997) such a battery of test species should be:

1. Representative of the ecosystem to protect, so including organisms having different life-histories, representing different functional groups, different taxonomic groups and different routes of exposure;

2. Representative of responses relevant for the protection of populations and communities; and

3. Uniform, so all tests in a battery should be applicable to the same test media and applying to the same test conditions, e.g. the same range of pH values.

Representation of environmental compartments

Each environmental compartment, water, air, soil and sediment, requires its specific set of test organisms. The most commonly applied test organisms are daphnids (*Daphnia magna*) for water, chironomids (*Chironomus riparius*) for sediments and earthworms (*Eisenia fetida*) for soil. For air, in the field of inhalation toxicology, humans and rodents are actually the most studied organism. In

ecotoxicology, air testing is mostly restricted to plants, concerning studies on toxic gasses. Besides the most commonly applied organisms, there is a long list of other standard test organisms for which test protocols are available (Table 1; OECD site).

Table 1. *Non-exhaustive list of standard ecotoxicity test species.*

Environmental compartment(s)	Organism group	Test species
Water	Plant	Myriophyllum spicatum
Water	Plant	Lemna
Water	Algae	Species of choice
Water	Cyanobacteria	Species of choice
Water	Fish	Danio rerio
Water	Fish	Oryzias latipes
Water	Amphibian	Xenopus laevis
Water	Insect	Chironomus riparius
Water	Crustacean	Daphnia magna
Water	Snail	Lymnaea stagnalis
Water	Snail	Potamopyrgus antipodarum
Water-sediment	Plant	Myriophyllum spicatum
Water-sediment	Insect	Chironomus riparius
Water-sediment	Oligochaete worm	Lumbriculus variegatus
Sediment	Anaerobic bacteria	Sewage sludge
Soil	Plant	Species of choice
Soil	Oligochaete worm	Eisenia fetida or E. andrei
Soil	Oligochaete worm	Enchytraeus albidus or E. crypticus
Soil	Collembolan	Folsomia candida or F. fimetaria
Soil	Mite	

		Hypoaspis (Geolaelaps) aculeifer
Soil	Microorganisms	Natural microbial community
Dung	Insect	Scathophaga stercoraria
Dung	Insect	Musca autumnalis
Air-soil	Plant	Species of choice
Terrestrial	Bird	Species of choice
Terrestrial	Insect	Apis mellifera
Terrestrial	Insect	Bombus terrestris/B. impatiens
Terrestrial	Insect	Aphidius rhopalosiphi
Terrestrial	Mite	Typhlodromus pyri

Non-standard test organisms

The use of standard test organisms in standard ecotoxicity tests performed according to internationally accepted protocols strongly reduces the uncertainties in ecotoxicity testing. Yet, there are good reasons for deviating from these protocols. The species in Table 1 are listed according to their corresponding environmental compartment, but ignores differences between ecosystems and habitats. Soils may differ extensively in composition, depending on e.g. the sand, clay or silt content, and properties, e.g. pH and water content, each harbouring different species. Likewise, stagnant and current water have few species in common. This implies that based on ecological arguments there may be good reasons to select non-standard test organisms. Effects of compounds in streams can be better estimated with riverine insects rather than with the stagnant water inhabiting daphnids, while the compost worm *Eisenia fetida* is not necessarily the most appropriate species for sandy soils. The list of non-standard test organisms is of course endless, but if the methods are well documented in the open literature, there are no limitations to employ these alternative species. They do involve, however, experimental challenges, since non-standard test organisms may be hard to culture and to maintain under laboratory conditions and no protocols are available for the ecotoxicity test. Thus increasing the ecological relevance of ecotoxicity tests also increases the logistical and experimental constraints (see chapter 6 on Risk assessment).

Increasing the number of test species

The vast majority of toxicity tests is performed with a single test species, resulting in large margins of uncertainty concerning the hazardousness of compounds. To reduce these uncertainties and to increase ecological relevance it is advised to incorporate more test species belonging to different trophic levels, for water e.g. algae, daphnids and fish. For deriving environmental quality standards from Species Sensitivity Distributions (see section on SSDs) toxicity data is required for minimal eight species belonging to different taxonomical groups. This obviously causes tension between the scientific requirements and the available financial resources.

References

OECD site. https://www.oecd-ilibrary.org/enviro...stems_20745761.

Van Gestel, C.A.M., Léon, C.D., Van Straalen, N.M. (1997). Evaluation of soil fauna ecotoxicity tests regarding their use in risk assessment. In: Tarradellas, J., Bitton, G., Rossel, D. (Eds). *Soil Ecotoxicology*. CRC Press, Inc., Boca Raton: 291-317.

4.3.4. Question 1

Name the requirements for suitable laboratory ecotoxicity test organisms.

4.3.4. Question 2

List the most commonly used standard test organisms per environmental compartment.

4.3.4. Question 3

Argue 1] the need for more than one test species, and 2] the need for non-standard test organisms.

4.3.5. Selection of test organisms - Eco plants

Author: J. Arie Vonk

Reviewers: Michiel Kraak, Gertie Arts, Sergi Sabater

Learning objectives:

You should be able to

- name the requirements for suitable laboratory ecotoxicity tests for primary producers
- list the most commonly used primary producers and endpoints in standardized ecotoxicity tests
- argue the need for selecting primary producers from different environmental compartments as test organisms for ecotoxicity tests

Key words: Test organism, standardized laboratory ecotoxicity test, primary producers, algae, plants, environmental compartment, photosynthesis, growth

Introduction

Photo-autotrophic primary producers use chlorophyll to convert CO_2 and H_2O into organic matter through photosynthesis under (sun)light. These primary producers are the basis of the food web and form an essential component of ecosystems. Besides serving as a food source, multicellular photo-autotrophs also form habitat for other primary producers (epiphytes) and many fauna species. Primary producers are a very diverse group, ranging from tiny unicellular pico-plankton up to gigantic trees. For standardized ecotoxicity tests, primary producers are represented by (micro)algae, aquatic macrophytes and terrestrial plants. Since herbicides are the largest group of pesticides used globally to maintain high crop production in agriculture, it is important to assess their impact on primary producers (Wang & Freemark, 1995). However, concerning testing intensity, primary producers are understudied in comparison to animals.

Standardized laboratory ecotoxicity tests with primary producers require good control over test conditions, standardized endpoints (Arts et al., 2008; see the Section on Endpoints) and growth in the controls (i.e. doubling of cell counts, length and/or biomass within the experimental period). Since the metabolism of primary producers is strongly influenced by light conditions, availability of water and inorganic carbon (CO_2 and/or HCO_3^- and CO_3^{2-}), temperature and dissolved nutrient concentrations, all these conditions should be monitored closely. The general criteria for selection of test organisms are described in the previous chapter (see the section on the Selection of ecotoxicity test organisms). For primary producers, the choice is mainly based on the available test guidelines, test species and the environmental compartment of concern.

Standardized ecotoxicity testing with primary producers

There are a number of ecotoxicity tests with a variety of primary producers standardized by different organizations including the OECD and the USEPA (Table 1). Characteristic for most primary producers is that they are growing in more than one

environmental compartment (soil/sediment; water; air). As a result of this, toxicant uptake for these photo-autotrophs might be diverse, depending on the chemical and the compartment where exposure occurs (air, water, sediment/soil).

For both marine and freshwater ecosystems, standardized ecotoxicity tests are available for microalgae (unicellular micro-organisms sometimes forming larger colonies) including the prokaryotic Cyanobacteria (blue-green algae) and the eukaryotic Chlorophyta (green algae) and Bacillariophyceae (diatoms). Macrophytes (macroalgae and aquatic plants) are multicellular organisms, the latter consisting of differentiated tissues, with a number of species included in standardized ecotoxicity tests. While macroalgae grow in the water compartment only, aquatic plants are divided into groups related to their growth form (emergent; free-floating; submerged and sediment-rooted; floating and sediment-rooted) and can extend from the sediment (roots and root-stocks) through the water into the air. Both macroalgae and aquatic plants contain a wide range of taxa and are present in both marine and freshwater ecosystems.

Terrestrial higher plants are very diverse, ranging from small grasses to large trees. Plants included in standardized ecotoxicity tests consist of crop and non-crop species. An important distinction in terrestrial plants is reflected in dicots and monocots, since both groups differ in their metabolic pathways and might reflect a difference in sensitivity to contaminants.

Table 1. *Open source standard guidelines for testing the effect of compounds on primary producers. All tests are performed in (micro)cosms except marked with **

Primary producer	Species	Compartment	Test number	Organisation
Microalgae & cyano-bacteria	various species	Freshwater	201	OECD 2011
	Anabaena flos-aque	Freshwater	850.4550	USEPA 2012
	Pseudokirchneriella subcapitata, Skeletonema costatum	Freshwater, Marine water	850.4500	USEPA 2012
Floating macrophytes	*Lemna* spp.	Freshwater	221	OECD 2006
	Lemna spp.	Freshwater	850.4400	USEPA 2012
Submerged macrophytes	Myriophyllum spicatum	Freshwater	238	OECD 2014
	Myriophyllum spicatum	Sediment (Freshwater)	239	OECD 2014
Aquatic plants*	not specified	Freshwater	850.4450	USEPA 2012
Terrestrial plants	wide variety of species	Air	227	OECD 2006
	wide variety of species	Air	850.4150	USEPA 2012
	wide variety of species (crops and non-crops)	Soil & Air	850.4230	USEPA 2012
	legumes and rhizobium symbiont	Soil & Air	850.4600	USEPA 2012
	wide variety of species (crops and non-crops)	Soil	208	OECD 2006
	wide variety of species (crops and non-crops)	Soil	850.4100	USEPA 2012
	various crop species	Soil & Air	850.4800	USEPA 2012
Terrestrial plants*	not specified	Terrestrial	850.4300	USEPA 2012

Representation of environmental compartments

Since primary producers can take up many compounds directly by cells and thalli (algae) or by their leaves, stems, roots and rhizomes (plants), different environmental compartments need to be included in ecotoxicity testing depending on the chemical characteristics of the contaminants. Moreover, the chemical characteristics of the compound under consideration determine if and how the compound might enter the primary producers and how it is transported through organisms.

For all aquatic primary producers, exposure through the water phase is relevant. Air exposure occurs in the emergent and floating aquatic plants, while rooting plants and algae with rhizoids might be exposed through sediment. Sediment exposure introduces

additional challenges for standardized testing conditions, since changes in redox conditions and organic matter content of sediments can alter the behavior of compounds in this compartment.

All terrestrial plants are exposed through air, soil and water (soil moisture, rain, irrigation). Air exposure and water deposition (rain or spraying) directly exposes aboveground parts of terrestrial plants, while belowground plant parts and seeds are exposed through soil and soil moisture. Soil exposure introduces additional challenges for standardized testing conditions, since changes in water or sediment organic matter content of soils can alter the behavior of compounds in this compartment.

Test endpoints

Bioaccumulation after uptake and translocation to specific cell organelles or plant tissue can result in incorporation of compounds in primary producers. This has been observed for heavy metals, pesticides and other organic chemicals. The accumulated compounds in primary producers can then enter the food chain and be transferred to higher trophic levels (see the section on Biomagnification). Although concentrations in primary producers are indicative of the presence of bioavailable compounds, these concentrations do not necessarily imply adverse effects on these organisms. Bioaccumulation measurements can therefore be best combined with one or more of the following endpoint assessments.

Photosynthesis is the most essential metabolic pathway for primary producers. The mode of action of many herbicides is therefore photosynthesis inhibition, whereby different metabolic steps can be targeted (see the section on Herbicide toxicity). This endpoint is relevant for assessing acute effects on the chlorophyll electron transport using Pulse-Amplitude-Modulation (PAM) fluorometry or as a measure of oxygen or carbon production by primary producers.

Growth represents the accumulation of biomass (microalgae) or mass (multicellular primary producers). Growth inhibition is the most important endpoint in test with primary producers since this endpoint integrates responses of a wide range of metabolic effects into a whole organism or a population response of primary producers. However, it takes longer to assess, especially for larger primary producers. Cell counts, increase in size over time for either leaves, roots, or whole organisms, and (bio)mass (fresh weight and dry weight) are the growth endpoints mostly used.

Seedling emergence reflects the germination and early development of seedlings into plants. This endpoint is especially relevant for perennial and biannual plants depending on seed dispersal and successful germination to maintain healthy populations.

Other endpoints include elongation of different plant parts (e.g. roots), necrosis of leaves, or disturbances in plant-microbial symbiont relationships.

Current limitations and challenges for using primary producers in ecotoxicity tests

For terrestrial vascular plants, many crop and non-crop species can be used in standardized tests, however, for other environmental compartments (aquatic and marine) few species are available in standardized test guidelines. Also not all environmental compartments are currently covered by standardized tests for primary producers. In general, there are limited tests for aquatic sediments and there is a total lack of tests for marine sediments. Finally, not all major groups of primary producers are represented in standardized toxicity tests, for example mosses and some major groups of algae are absent.

Challenges to improve ecotoxicity tests with plants would be to include more sensitive and early response endpoints. For soil and sediment exposure of plants to contaminants, development of endpoints related to root morphology and root metabolism could provide insights into early impact of substances to exposed plant parts. Also the development of ecotoxicogenomic endpoints (e.g. metabolomics) (see the section on Metabolomics) in the field of plant toxicity tests would enable us to determine effects on a wider range of plant metabolic pathways.

References

Arts, G.H.P., Belgers, J.D.M., Hoekzema, C.H., Thissen, J.T.N.M. (2008). Sensitivity of submersed freshwater macrophytes and endpoints in laboratory toxicity tests. Environmental Pollution 153, 199-206.

Wang, W.C., Freemark, K. (1995) The use of plants for environmental monitoring and assessment. Ecotoxicology and Environmental Safety 30: 289-301.

4.3.5. Question 1

Which conditions need to be controlled carefully in laboratory ecotoxicity tests with primary producers?

4.3.5. Question 2

List the different groups of primary producers used in standardized tests for each environmental compartment.

4.3.5. Question 3

Argue why testing of primary producers is relevant in relation to [A] environmental exposure of ecosystems to pesticides and [B] the role of primary producers in ecosystems

4.3.6. Selection of test organisms - Microorganisms

Author: Patrick van Beelen

Reviewers: Kees van Gestel, Erland Bååth, Maria Niklinska

Learning objectives:

You should be able to

- describe the vital role of microorganisms in ecosystems.
- explain the difference between toxicity tests for protecting biodiversity and for protecting ecosystem services.
- explain why short-term microbial tests can be more sensitive than long-term ones.

Keywords: microorganisms, processes, nitrogen conversion, test methods

The importance of microorganisms

Most organisms are microorganisms, which means they are generally too small to see with the naked eye. Nevertheless, microorganisms affect almost all aspects of our lives. Viruses are the smallest of microorganisms, the prokaryotic bacteria and archaea are bigger (in the micrometer range), and the sizes of eukaryotic microorganisms range from three to hundred micrometers. The microscopic eukaryotes have larger cells with a nucleus and come in different shapes like green algae, protists and fungi.

Cyanobacteria and eukaryotic algae perform photosynthesis in the oceans, seas, brackish and freshwater ecosystems. They fix carbon dioxide into biomass and form the basis of the largest aquatic ecosystems. Bacteria and fungi degrade complex organic molecules into carbon dioxide and minerals, which are needed for plant growth.

Plants often live in symbiosis with specialized microorganisms on their roots, which facilitate their growth by enhancing uptake of water and nutrients, speeding up plant growth. Invertebrate and vertebrate animals, including humans, have bacteria and other microorganisms in their intestines to facilitate the digestion of food. Cows for example cannot digest grass without the microorganisms in their rumen. Also, termites would not be able to digest lignin, a hard to digest wood polymer, without the aid of gut fungi. Leaf cutter ants transport leaves into their nest to feed the fungi which they depend on. Also, humans consume many foodstuffs with yeasts, fungi or bacteria for preservation of the food and a pleasant taste. Beer, wine, cheese, yogurt, sauerkraut, vinegar, bread, tempeh, sausage and may other foodstuffs need the right type of microorganisms to be palatable. Having the right type of microorganisms is also vital for human health. Human mother's milk contains oligosaccharides, which are indigestible for the newborn child. These serve as a major food source for the intestinal bacteria in the baby, which reduce the risk of dangerous infections.

This shows that the interaction between specific microorganisms and higher organisms are often highly specific. Marine viruses are very abundant and can limit algal blooms promoting a more diverse marine phytoplankton. Pathogenic viruses, bacteria, fungi and protists enhance the biodiversity of plants and animals by the following mechanism: The densest populations are more susceptible to diseases since the transmission of the disease becomes more frequent. When the most abundant species become less frequent, there is more room for the other species and biodiversity is enhanced. In agriculture, this enhanced biodiversity is unwanted since the livestock and crop are the most abundant species. That is why disease control becomes more important in high intensity livestock farming and in large monocultures of crops. Microorganisms are at the base of all ecosystems and are vital for human health and the environment.

The microbiological society has a nice video explaining why microbiology matters.

Protection goals

The functioning of natural ecosystems on earth is threatened by many factors, such as habitat loss, habitat fragmentation, global warming, species extinction, over fertilization, acidification and pollution. Natural and man-made chemicals can exhibit toxic effects on the different organisms in natural ecosystems. Toxic chemicals released in the environment may have negative effects on biodiversity or microbial processes. In the ecosystem strongly affected by such changes, the abundance of different species could be smaller. The loss of biodiversity of the species in a specific ecosystem can be used as a measure for the degradation of the ecosystem. Humans benefit from the presence of properly functioning ecosystems. These benefits can be quantified as ecosystem services. Microbial processes contribute heavily to many ecosystem services. Groundwater for example, is often a suitable source of drinking water since microorganisms have removed pollutants and pathogens from the infiltrating water. See Section on Ecosystem services and protection goals.

Environmental toxicity tests

Most environmental toxicity tests are single species tests. Such tests typically determine toxicity of a chemical to a specific biological species like for example the bioluminescence by the *Allivibrio fisheri* bacteria in the Microtox test or the growth inhibition test on freshwater algae and cyanobacteria (see Section on Selection of test organisms - Eco plants). These tests are relatively simple using a specific toxic chemical on a specific biological species in an optimal setting. The OECD guidelines for the testing of chemicals, section 2, effects on biotic systems gives a list of standard tests. Table 1 lists different tests with microorganisms standardized by the Organization for Economic Cooperation and Development (OECD).

Table 1. Generally accepted environmental toxicity tests using microorganisms, standardized by the Organization for Economic Cooperation and Development (OECD).

OECD test No	Title	Medium	Test type
201	Freshwater algae and cyanobacteria, growth inhibition test (chapter reference)	Aquatic	Single species
209	Activated sludge, respiration inhibition test	Sediment	Process
224 (draft guideline)	Determination of the inhibition of the activity of anaerobic bacteria	Sediment	Process
217	Soil microorganisms: carbon transformation test	Soil	Process
216	Soil microorganisms: nitrogen transformation test	Soil	Process

The outcome of these tests can be summarized as EC_{10} values (see Section on Concentration-response relationships), which can be used in risk assessment (see Sections on Predictive risk assessment approaches and tools and on Diagnostic risk assessment approaches and tools) Basically, there are three types of tests. Single species tests, community tests and tests using microbial processes.

Single species tests

The ecological relevance of a single species test can be a matter of debate. In most cases it is not practical to work with ecologically relevant species since these can be hard to maintain under laboratory conditions. Each ecosystem will also have its own ecologically relevant species, which would require an extremely large battery of different test species and tests, which are difficult to perform in a reproducible way. As a solution to these problems, the test species are assumed to exhibit similar sensitivity for toxicants as the ecological relevant species. This assumption was confirmed in a number of cases. If the sensitivity distribution

of a given toxicant for a number of test species would be similar to the sensitivity distribution of the relevant species in a specific ecosystem, one could use a statistic method to estimate a safe concentration for most of the species.

Toxicity tests with short incubation times are often disputed since it takes time for toxicants to accumulate in the test animals. This is not a problem in microbial toxicity tests since the small size of the test organisms allows a rapid equilibrium of the concentrations of the toxicant in the water and in the test organism. On the contrary, long incubation times under conditions that promote growth, can lead to the occurrence of resistant mutants, which will decrease the apparent sensitivity of the test organism. This selection and growth of resistant mutants cannot, however, be regarded as a positive thing since these mutants are different from the parent strain and might also have different ecological properties. In fact, the selection of antibiotic resistant microorganisms in the environment is considered to be a problem since these might transfer to pathogenic (disease promoting) microorganisms which gives problems for patients treated with antibiotics.

The OECD test no 201, which uses freshwater algae and cyanobacteria, is a well-known and sensitive single species microbial ecotoxicity test. These are explained in more detail in the Section on Selection of test organisms - Eco plants.

Community tests

Microorganisms have a very wide range of metabolic diversity. This makes it more difficult to extrapolate from a single species test to all possible microbial species including fungi, protists, bacteria, archaea and viruses. One solution is to test a multitude of species (a whole community) exposed in a single toxicity experiment, it becomes more difficult to attribute the decline or increase of species to toxic effects. The rise and decline of species can also be caused by other factors, including species interactions. The method of Pollution-induced community tolerance is used for the detection of toxic effects on communities. Organisms survive in polluted environments only when they can tolerate toxic chemical concentrations in their habitat. During exposure to pollution the sensitive species become extinct and tolerant species take over their place and role in the ecosystem (Figure 1). This takeover can be monitored by very simple toxicity tests using a part of the community extracted from the environment. Some tests use the incorporation of building blocks for DNA (thymidine) and protein (leucine). Other tests use different substrates for microbial growth. The observation that this part of the community becomes more tolerant as measured by these simple toxicity tests reveals that the pollutant really affects the microbial community. This is especially helpful when complex and diverse environments like biofilms, sediments and soils are studied.

Tests using microbial processes

The protection of ecosystem services is fundamentally different from the protection of biodiversity. When one wants to protect biodiversity all species are equally important and are worth protecting. When one wants to protect ecosystem services only the species that perform the process have to be protected. Many contributing species can be intoxicated without having much impact on the process. An example is nitrogen transformation, which is tested by measuring the conversion of ammonium into nitrite and nitrate (see box).

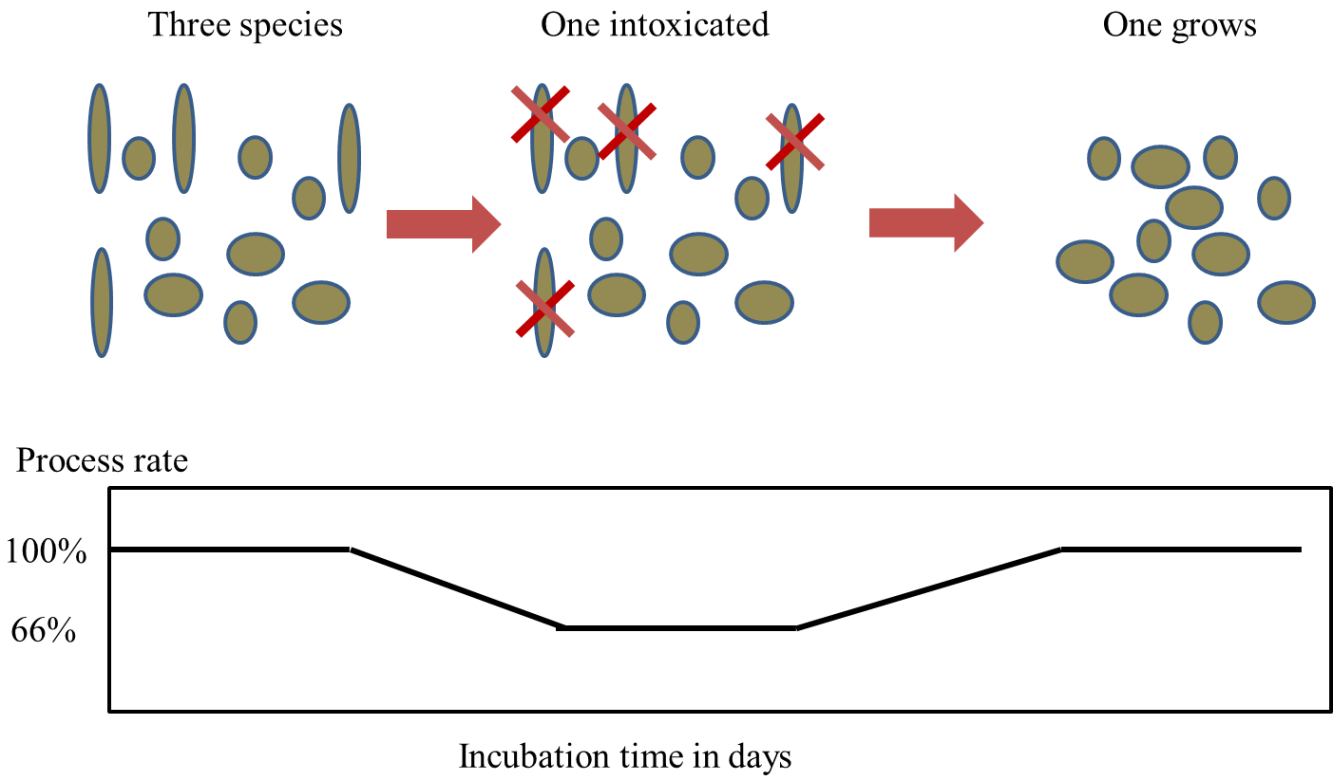


Figure 1. The effect of growth on an intoxicated process performed by different species of microorganisms. The intoxication of some species may temporarily decrease process rate, but due to growth of the tolerant species this effect soon disappears and process rate is restored. Source: Patrick van Beelen.

The inactivation of the most sensitive species can be compensated by the prolonged activity or growth of less sensitive species. The test design of microbial process tests aims to protect the process and not the contributing species. Consequently, the process tests

from Table 1 seldom play a decisive role in reducing the maximum tolerable concentration of a chemical. Reason is that the single species toxicity tests generally are more sensitive since they use a specific biological species as test organism instead of a process.

Box: Nitrogen transformation test

The OECD test no. 216 Soil Microorganisms: Nitrogen Transformation Test is a very well-known toxicity test using the soil process of nitrogen transformation. The test for non-agrochemicals is designed to detect persistent adverse effects of a toxicant on the process of nitrogen transformation in soils. Powdered clover meal contains nitrogen mainly in the form of proteins which can be degraded and oxidized to produce nitrate. Soil is amended with clover meal and treated with different concentrations of a toxicant. The soil provides both the test organisms and the test medium. A sandy soil with a low organic carbon content is used to minimize sorption of the toxicant to the soil. Sorption can decrease the toxicity of a toxicant in soil. According to the guideline, the soil microorganisms should not be exposed to fertilizers, crop protection products, biological materials or accidental contaminations for at least three months before the soil is sampled. In addition, the soil microorganisms should at least form 1% of the soil organic carbon. This indicates that the microorganisms are still alive. The soil is incubated with clover meal and the toxicant under favorable growth conditions (optimal temperature, moisture) for the microorganisms. The quantities of nitrate formed are measured after 7 and 28 days of incubation. This allows for the growth of microorganisms resistant to the toxicant during the test, which can make the longer incubation time less sensitive. The nitrogen in the proteins of clover meal will be converted to ammonia by general degradation processes. The conversion of clover meal to ammonia can be performed by a multitude of species and is therefore not very sensitive to inhibition by toxic compounds.

$$clovermeal \rightarrow proteins \rightarrow aminoacids \rightarrow NH_4^+ \rightarrow NO_2^- \rightarrow NO_3^-$$

The conversion of ammonia to nitrate generally is performed in two steps. First, ammonia oxidizing bacteria or archaea, oxidize ammonia into nitrite. Second, nitrite is oxidized by nitrite oxidizing bacteria into nitrate. These latter two steps are generally much slower than ammonium production, since they require specialized microorganisms. These specialized microorganisms also have a lower growth rate than the common microorganisms involved in the general degradation of proteins into amino acids. This makes the nitrogen transformation test much more sensitive compared to the carbon transformation test, which uses more common microorganisms. Under the optimal conditions in the nitrogen transformation test some minor ammonia or nitrite oxidizing species might seem unimportant since they do not contribute much to the overall process. Nevertheless these minor species can become of major importance under less optimal conditions. Under acid conditions for example, only the archaea oxidize ammonia into nitrite while the ammonia oxidizing bacteria become inhibited. The nitrogen transformation test has a minimum duration of 28 days at 20°C under optimal moisture conditions, but can be prolonged to 100 days. Shorter incubation times would make the test more sensitive.

4.3.6. Question 1

What is the disadvantage of growth during a toxicity test using a microbial process?

4.3.6. Question 2

Mention the intermediates during the degradation of clover meal to nitrate.

4.3.6. Question 3

Why are shorter microbial tests often more sensitive than longer ones?

4.3.6. Question 4

When one microbial species in the environment is replaced by another one, can it have an effect on animals or plants?

4.3.7. Selection of test organisms - Birds

Author: Annegaaike Leopold

Reviewers: Nico van den Brink, Kees van Gestel, Peter Edwards

Learning objectives:

You should be able to

- Understand and argue why birds are an important model in ecotoxicology;
- understand and argue the objective of avian toxicity testing performed for regulatory purposes;

- list the most commonly used avian species;
- list the endpoints used in avian toxicity tests;
- name examples of how uncertainty in assessing the risk of chemicals to birds can be reduced.

Keywords: birds, risk assessment, habitats, acute, reproduction.

Introduction

Birds are seen as important models in ecotoxicology for a number of reasons:

- they are a diverse, abundant and widespread order inhabiting many human altered habitats like agriculture;
- they have physiological features that make them different from other vertebrate classes that may affect their sensitivity to chemical exposure;
- they play a specific role ecologically and fulfill essential roles in ecosystems (e.g. in seed dispersal, as biological control agents through eating insects, and removal of carcasses e.g. by vultures);
- protection goals are frequently focused on iconic species that appeal to the public.

A few specific physiological features will be discussed here. Birds are oviparous, laying eggs with hard shells. This leads to concentrated exposure (as opposed to exposure via the bloodstream as in most other vertrebrate species) to maternally transferred material, and where relevant, its metabolites. It also means that offspring receive a single supply of nutrients (and not a continuous supply through the blood stream). This makes birds sensitive to contaminants in a different way than non-oviparous vertebrates, since the embryos develop without physiological maternal interference. The bird embryo starts to regulate its own hormone homeostasis early on in its development in contrast to mammalian embryos. As a result contaminants deposited in the egg by the female bird may cause disturbance of the regulation of these embryonic processes (Murk et al., 1996). Birds have a higher body temperature (40.6 ºC) and a relatively high metabolic rate, which can impact their response to chemicals. As chicks, birds generally have a rapid growth rate, compared to many vertebrate species. Chicks of precocial (or nidifugous) species leave the nest upon hatching and, while they may follow the parents around, they are fully feathered and feed independently. They typically need a few months to grow to full size. Altricial species are naked, blind and helpless at hatch and require parental care until they fledge the nest. They often grow faster - passerines (such as swallows) can reach full size and fledge 14 days after hatching. Many bird species migrate seasonally over long distances and adaptation to this, changes their physiology and biochemical processes. Internal concentrations of organic contaminants, for example, may increase significantly due to the use of lipids stores during migration, while changes in biochemistry may increase the sensitivity of birds to the chemical.

Birds function as good biological indicators of environmental quality largely because of their position in the foodchain and habitat dependence. Protection goals are frequently focused on iconic species, for example the Atlantic puffin, the European turtle dove and the common barn owl (Birdlife International, 2018).

It was recognized early on that exposure of birds to pesticides can take place through many routes of dietary exposure. Given their association with a wide range of habitats, exposure can take place by feeding on the crop itself, on weeds, or (treated) weed seeds, on ground dwelling or foliar dwelling invertebrates, by feeding on invertebrates in the soil, such as earthworms, by drinking water from contaminated streams or by feeding on fish living in contaminated streams (Figure 1, Brooks et al., 2017). Following the introduction of persistent and highly toxic synthetic pesticides in the 1950s and prior to safety regulations, use of many synthetic organic pesticides led to wildlife losses - of birds, fish, and other wildlife (Kendall and Lacher, 1994). As a result, national and international guidelines for assessing first acute and subacute effects of pesticides on birds were developed in the 1970s. In the early 1980s tests were developed to study long-term or reproductive effects of pesticides. Current bird testing guidelines focused primarily on active ingredients used in plant protection products, veterinary medicines and biocides. In Europe the industrial chemicals regulation REACH only requires information on long-term or reproductive toxicity for substances manufactured or imported in quantities of at least 1000 tonnes per annum. These data may be needed to assess the risks of secondary poisoning by a substance that is likely to bioaccumulate and does not degrade rapidly. Secondary poisoning may occur, for example when raptors consume contaminated fish. In the United States no bird tests are required under the industrial chemicals legislation.

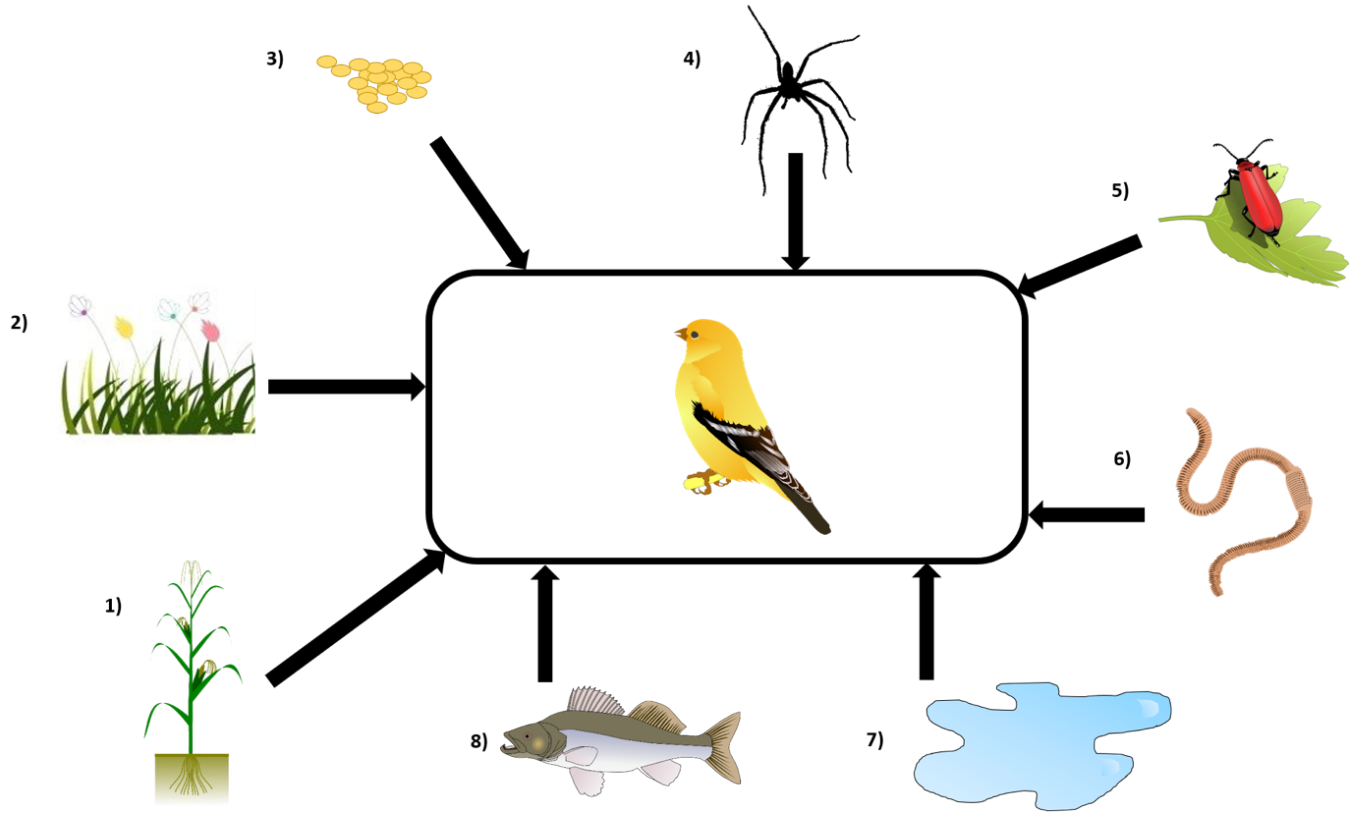


Figure 1. Potential routes of dietary exposure for birds feeding in agricultural fields sprayed with a crop protection product (pesticide). Most of the pesticide will land up in the treated crop area, but some of it may land in neighbouring surface water. Exposure to birds can therefore take place through many routes: by feeding on the crop itself (1), on weeds (2), or weed seeds (3), on ground-dwelling (**4**) or foliar-dwelling (**5**) invertebrates. Birds may also feed on earthworms living in the treated soil (**6**). Exposure may also occur by drinking from contaminated puddles within the treated crop area (**7**) or birds may feed on fish living in neighbouring contaminated surface waters (**8**). Based on Brooks et al. (2017).

The objective of performing avian toxicity tests is to inform an avian effects assessment (Hart et al., 2001) in order to:

- provide scientifically sound information on the type, size, frequency and pattern over time of effects expected from defined exposures of birds to chemicals.
- reduce uncertainty about potential effects of chemicals on birds.
- provide information in a form suitable for use in risk assessment.
- provide this information in a way that makes efficient use of resources and avoids unnecessary use and suffering of animals.

Bird species used in toxicity testing

Selection of bird species for toxicity testing occurs primarily on the basis of their ecological relevance, their availability and ability to adjust to laboratory conditions for breeding and testing. This means most test species have been domesticated over many years. They should have been shown to be relatively sensitive to chemicals through previous experience or published literature and ideally have available historical control data.

The bird species most commonly used in toxicity testing have all been domesticated:

- the waterfowl species mallard duck (*Anas platyrynchos)* is in the mid range of sensitivity to chemicals, an omnivorous feeder, abundant in many parts of the world, a precocial species; raised commercially and test birds show wild type plumage;
- the ground dwelling game species bobwhite quail (*Colinus virginianus)* is common in the USA and is of similar in sensitivity to mallards; feeds primarily on seeds and invertebrates; a precocial species; raised commercially and test birds show wild type plumage;
- the ground dwelling species Japanese quail (*Coturnix coturnix japonica*) occurs naturally in East Asia; feeds on plants material and terrestrial invertebrates. Domesticated to a far great extent than mallard or bobwhite quail and birds raised commercially

(for eggs or for meat) are further removed genetically from the wild type. This species is unique in that the young of the year mature and breed (themselves) within 12 months;
- the passerine, altricial species zebra finch *(Taeniopygia guttata*) occurs naturally in Australia and Indonesia; they eat seeds; are kept and sold as pets; are not far removed from wild type;
- the budgerigar (*Melopsittacus undulates*) is also altricial); occurs naturally in Australia; eats seeds eating; is bred in captivity and kept and sold as pets.

Other species of birds are sometimes used for specific, often tailor-designed studies. These species include:

- the canary (*Serinus canaria domestica).*
- the rock pigeon (*Columba livia)*
- the house sparrow (*Passer domesticus),*
- red-winged blackbird (*Agelaius phoeniceus)* - US only
- the ring-necked pheasant, (*Phasianus colchicus),*
- the grey partridge (*Perdix perdix)*

Most common avian toxicity tests:

Table 1 provides an overview of all the avian toxicity tests that have been developed over the past approximately 40 years, the most commonly used guidelines, the recommended species, the endpoints recorded in each of these tests, the typical age of birds at the start of the test, the test duration and the length of exposure.

Table 1: Most common avian toxicity tests with their recommended species and key characteristics.

Avian toxicity test	Guideline	Recommended species	Endpoints	Age at start of test	Length of study	Length of exposure
Acute oral gavage-sequential testing - average 26 birds	OECD 223	bobwhite quail, Japanese quail, zebra finch, budgerigar	mortality, clinical signs, body weight, food consumption, gross necropsy	Young birds not yet mated, at least 16 weeks old at start of test.	At least 14 days	Single oral dose at beginning of test
Acute oral gavage - 60 bird design	USEPA OCSPP 850.2100	Bobwhite quail single passerine species recommended.	See above	Young birds not yet mated, at least 16 weeks old at start of test.	14 days	Single oral dose at beginning of test
Sub-acute dietary toxicity *	OCSPP 850.2200	Bobwhite quail, mallard	See above	Mallard: 5 days old Bobwhite quail: 10-14 days old	8 days	5 days
One-generation reproduction	OECD 206 OCSPP 850.2200	Bobwhite quail, mallard, Japanese quail**	Adult body weight,food consumption, egg production, fertility, embryo survivial, hatchrate, chick survival.	Approaching first breeding season: Mallard (6 to 12 months old) Bobwhite quail 20 -24 weeks	20 - 22 weeks	10 weeks
Avoidance testing (pen trials)	OECD Report	As closely related to species at risk as possible; eg: sparrow rock dove, pheasant, grey partridge	Food intake, mortality, Sublethal effects	Young adults if possible (depends on study design)	One to several days, depending on the study design.	One to several days, depending on the study design.
Two-generation endocrine disruptor test	OCSPP 890.2100	Japanese quail	In addition to endpoints listed for one-generation study: male sexual behaviour, biochemical, histological, and morphological endpoints	4 weeks post hatch	38 weeks	8 weeks - adult (F0) generation +14 weeks F1 generation.
Field studies to refine food residues in	Appendix N of the EFSA Bird and Mammal Guidance.	Depends on the species at risk in the are of pesticide use.	Depends on the study design developed.	Uncontrollable in a field study	Depends on the study design developed.	Depends on the study design developed.

higher tier Bird risk assessments.						

* This study is hardly every asked for anymore.

** Only in OECD Guideline

Acute toxicity testing

To assess the short-term risk to birds, acute toxicity tests must be performed for all pesticides (the active ingredient thereof) to which birds are likely to be exposed, resulting in an LD_{50} (mg/kg body/day) (see section on Concentration-response relationships). The acute oral toxicity test involves gavage or capsule dosing at the start of the study (Figure 2). Care must be taken when dosing birds by oral gavage. Some species can readily regurgitate leading to uncertainty in the the dose given. These include mallard duck, pigeons and some passerine species. Table 1 gives the birds species recommended in the OECD and USEPA guidelines, respectively. Gamebirds and passerines are a good combination to take account of phylogeny and a good starting point to better understand the distribution of species sensitivity.

The OECD guideline 223 uses on average 26 birds and is a sequential design (Edwards et al., 2017). Responses of birds to each stage of the test are combined to estimate and improve the estimate of the LD_{50} and slope. The testing can be stopped at any stage once the accuracy of the LD_{50} estimate meets the requirements for the risk assessment, hence using far fewer birds for compliance with the 3Rs (reduction, refinement and replacement). If toxicity is expected to be low, 5 birds are dosed at the limit dose of 2000 mg/kg (which is the highest acceptable dose to be given by oral gavage, from a humane point of view). If there is no mortality in the limit test after 14 days the study is complete and the LD_{50} >2000 mg/kg body weight. If there is mortality a single individual is treated at each of 4 different doses in Stage 1. With these results a working estimate of the LD_{50} is determined to select 10 further dose levels for a better estimate of the LD_{50} in Stage 2. If a slope is required a further Stage 3 is required using 10 more birds in a combination of doses selected on the basis of a provisional estimate of the slope.

The USEPA guideline is a single stage design preceeded by a range finding test (used only to set the concentrations for the main test). The LD_{50} test uses 60 birds (10 at each of five test concentrations and 10 birds in the control group). Despite the high numbers of birds used, the ability to estimate a slope is poor compared to OECD223 (the ability to calculate the LD_{50} is similar to the OECD 223 guideline).



Figure 2. Gavage dosing of a zebrafinch - Eurofins Agroscience Services, Easton MD, USA.

Dietary toxicity testing

For the medium-term risk assessment an avian dietary toxicity test was regularly performed in the past exposing juvenile (chicks) of bobwhite quail, Japanese quail or mallard to a treated diet. This test determines the median lethal concentration (LC_{50}) of a chemical in response to a 5-day dietary exposure. Given the scientific limitations and animal welfare concerns related to this test (EFSA, 2009) current European regulations recommend to only perform this test when it is expected that the LD_{50} value measured by the medium-term study will be lower than the acute LD_{50} i.e. if the chemical is cumulative in its effect.

Reproduction testing

One-generation reproduction tests in bobwhite quail and/or mallard are requested for the registration of all pesticides to which birds are likely to be exposed during the breeding season. Table 1 presents the two standard studies: OECD Test 206 and the US EPA OCSPP 850.2100 study. The substance to be tested is mixed into the diet from the start of the test. The birds are fed *ad libitum* for a recommended period of 10 weeks before they begin laying eggs in response to a change in photoperiod. The egg-laying period should last at least ten weeks. Endpoints include adult body weight, food consumption, macroscopic findings at necropsy and reproductive endpoints, with the number of 14-day old surviving chicks/ducklings as an overall endpoint.

The OECD guideline states that the Japanese quail (*Coturnix coturnix japonica*), is also acceptable.

Avoidance (or repellancy) testing

Avoidance behaviour by birds in the field could be seen as reducing the risk of exposure to a pesticide and therefore could be considered in the risk assessment. However, the occurrence of avoidance in the laboratory has a confounding effect on estimates of toxicity in dietary studies (LD_{50}). Avoidance tests thus far have greatest relevance in the risk assessment of seed treatments. A number of factors need to be taken into account including the feeding rate and dietary concentration which may determine whether avoidance or mortality is the outcome. The following comprehensive OECD report provides an overview of guideline development and research activities that have taken place to date under the OECD flag. Sometimes these studies are done as semi-field (or pen) studies.

Endocrine disruptor testing

Endocrine-disrupting substances can be defined as materials that cause effects on reproduction through the disruption of endocrine-mediated processes. If there is reason to suspect that a substance might have an endocrine effect in birds, a two-generation avian test design aimed specifically at the evaluation of endocrine effects could be performed. This test has been developed by the USEPA (OCSPP 890.210). The test has not, however, been accepted as an OECD test to date. It uses the Japanese quail as the preferred species. The main reasons that Japanese quail were selected for this test were: 1) Japanese quail is a precocial species as mentioned earlier. This means that at hatch Japanese quail chicks are much further in their sexual differentiation and development than chicks of altricial species would be. Hormonal processes occurring in Japanese quail in these early stages of development can be disturbed by chemicals maternally deposited in the egg (Ottinger and Dean, 2011). Conversely altricial species undergo these same sexual development stages post-hatch and can be exposed to chemicals in food that might impact these same hormonal processes. 2) as mentioned above, the young of the year mature and breed (themselves) within 12 months which makes the test more efficient that if one used bobwhite quail or mallard.

It is argued among avian toxicologists, that it is necessary to develop a zebra finch endocrine assay system, alongside the Japanese quail system, as this will allow a more systematic determination of differences between responses to EDC's in altricial and precocial species, there by allowing a better evaluation and subsequent risk assessment of potential endocrine effects in birds. Differences in parental care, nesting behaviour and territoriality are examples of aspects that could be incorporated in such an approach (Jones et al., 2013).

Field studies:

Field studies can be used to test for adverse effects on a range of species simultaneously, under conditions of actual exposure in the environment (Hart et al, 2001). The numbers of sites and control fields and methods (corpse searches, censusing and radiotracking) need careful consideration for optimal use of field studies in avian toxicology. The field site will define the species studied and it is

important to consider the relevance of that species in other locations. For further reading about techniques and methods to be used in avian field research Sutherland et al and Bibby et al. (2000) are recommended.

References

Bibby, C., Jones, M., Marsden, S. (2000). Expedition Field Techniques Bird Surveys. Birdlife International.

Birdlife International (2018). The Status of the World's Birds. https://www.birdlife.org/sites/default/files/attachments/BL_ReportENG_V11_spreads.pdf

Brooks, A.C., Fryer, M., Lawrence, A., Pascual, J., Sharp, R. (2017). Reflections on bird and mammal risk assessment for plant protection products in the European Union: Past, present and future. Environmental Toxicology and Chemistry 36, 565-575.

Edwards, P.J., Leopold, A., Beavers, J.B., Springer, T.A., Chapman, P., Maynard, S.K., Hubbard, P. (2017). More or less: Analysis of the performance of avian acute oral guideline OECD 223 from empirical data. Integrated Environmental Assessment and Management 13, 906-914.

Hart, A., Balluff, D., Barfknecht, R., Chapman, P.F., Hawkes, T., Joermann, G., Leopold, A., Luttik, R. (Eds.) (2001). Avian Effects Assessment: A Framework for Contaminants Studies. A report of a SETAC workshop on 'Harmonised Approaches to Avian Effects Assessment', held with the support of the OECD, in Woudschoten, The Netherlands, September 1999. A SETAC Book.

Jones, P.D., Hecker, M., Wiseman, S., Giesy, J.P. (2013). Birds. Chapter 10 In: Matthiessen, P. (Ed.) Endocrine Disrupters - Hazard Testing and Assessment Methods. Wiley & Sons.

Kendall, R.J., Lacher Jr, T.E. (Eds.) (1994). Wildlife Toxicology and Population Modelling - Integrated Studies of Agrochecosystems. Special Publication of SETAC.

Murk, A.J., Boudewijn, T.J., Meininger, P.L., Bosveld, A.T.C., Rossaert, G., Ysebaert, T., Meire, P., Dirksen, S. (1996). Effects of polyhalogenated aromatic hydrocarbons and related contaminants on common tern reproduction: Integration of biological, biochemical, and chemical data. Archives of Environmental Contamination and Toxicology 31, 128-140.

Ottinger, M.A., Dean, K. (2011). Neuroendocrine Impacts of Endocrine-Disrupting Chemicals in Birds: Life Stage and Species Sensitivities. Journal of Toxicology and Environmental Health, Part B: Critical Reviews. 26 July 2011.

Sutherland, W.J., Newton, I., Green, R.E. (Eds.) (2004). Biological Ecology and Conservation. A Handbook of Techniques. Oxford University Press

4.3.7. Question 1

Give three reasons why birds are an important model in ecotoxicology?

4.3.7. Question 2

What is the objective of avian toxicity testing?

4.3.7. Question 3

Which avian species are most commonly used and why?

4.3.7. Question 4

What is the difference between precocial and altricial species?

4.3.7. Question 5

Which endpoints are typically used in standardised avian toxicity tests?

4.3.7. Question 6

Name two examples of how one might reduce uncertainty in assessing risks of chemicals on birds.

4.3.8. In vitro toxicity testing

Author: Timo Hamers

Reviewer: Arno Gutleb

Learning goals

You should be able to:

- explain the difference between in vitro and in vivo bioassays
- describe the principle of a ligand binding assay, an enzyme inhibition assay, and a reporter gene bioassay
- explain the difference between primary cell cultures, finite cell lines, and continuous cell lines
- describe different levels in cell differentiation potency from totipotent to unipotent;
- indicate how in vitro cell cultures of differentiated cells can be obtained from embryonic stem cells and from induced pluripotent stem cells
- give examples of endpoints that can be measured in cell-based bioassays
- discuss in his own words a future perspective of in vitro toxicity testing

Keywords: ligand binding assay; enzyme inhibition assay; primary cell culture; cell line; stem cell; organ on a chip

Introduction

In vitro bioassays refer to testing methods making use of tissues, cells, or proteins. The term "in vitro" (meaning "in glass") refers to the test tubes or petri dishes made from glass that were traditionally used to perform these types of toxicity tests. Nowadays, in vitro bioassays are more often performed in plastic microtiter wells-plates containing multiple (6, 12, 24, 48, 96, 384, or 1536) test containers (called "wells") per plate (Figure 1). In vitro bioassays are usually performed to screen individual substances or samples for specific bioactive properties. As such, in vitro toxicology refers to the science of testing substances or samples for specific toxic properties using tissues, cells, or proteins.



Figure 1: Six different microtiter well-plates, consisting of multiple small-volume test containers. In clockwise direction starting from the LOWER left: 6-wells plate, 12-wells plate, 24-wells plate, 48-wells plate, 96 wells plate, 384 wells plate.

Most in vitro bioassays show a mechanism-specific response, which is for instance indicative of the inhibition of a specific enzyme or the activation of a specific molecular receptor. Moreover, in vitro bioassays are usually performed in small test volumes and have short test durations (usually incubation periods range from 15 minutes to 48 hours). As a consequence, multiple samples can be tested simultaneously in a single experiment and multiple experiments can be performed in a relatively short test period. This "medium-throughput" characteristic of in vitro bioassays can even be increased to high-throughput" if the time-limiting steps in the test procedure (e.g. sample preparation, cell culturing, pipetting, read-out) are further automated.

Toxicity tests making use of bacteria are also often performed in small volumes, allowing short test-durations and high-throughput. Still, such tests make use of intact organisms and should therefore strictly be considered as in vivo bioassays. This holds especially true if bacteria are used to study endpoints like survival or population growth. However, bacteria test systems studying specific toxic mechanisms, such as the Ames test used to screen substances for mutagenic properties (see section on Carcinogenicity and Genotoxicity), are often considered as in vitro bioassays, because of the similarity in test characteristics when compared to in vitro toxicity tests with cells derived from higher organisms.

Protein-based assays

The simplest form of an in vitro binding assay consists of a purified protein that is incubated with a potential toxic substance or sample. Purified proteins are usually obtained by isolation from an intact organism or from cultures of recombinant bacteria, which are genetically modified to express the protein of interest.

Ligand binding assays are used to determine if the test substance is capable of binding to the protein, thereby inhibiting the binding capacity of the natural (endogenous) ligand to that protein (see section on Protein Inactivation). Proteins of interest are for instance receptor proteins or transporter proteins. Ligand binding assays often make use of a natural ligand that has been labelled with a radioactive isotope. The protein is incubated with the labelled ligand in the presence of different concentrations of the test substance. If protein-binding by the test substance prevents ligand binding to the protein, the free ligand shows a concentration-dependent increase in radioactivity (See Figure 2). Consequently, the ligand-protein complex shows a concentration-dependent decrease in radioactivity. Alternatively, the natural ligand may be labelled with a fluorescent group. Binding of such a labelled ligand to the protein often causes an increase in fluorescence. Consequently, a decrease in fluorescence is observed if a test substance prevents ligand binding to the protein.

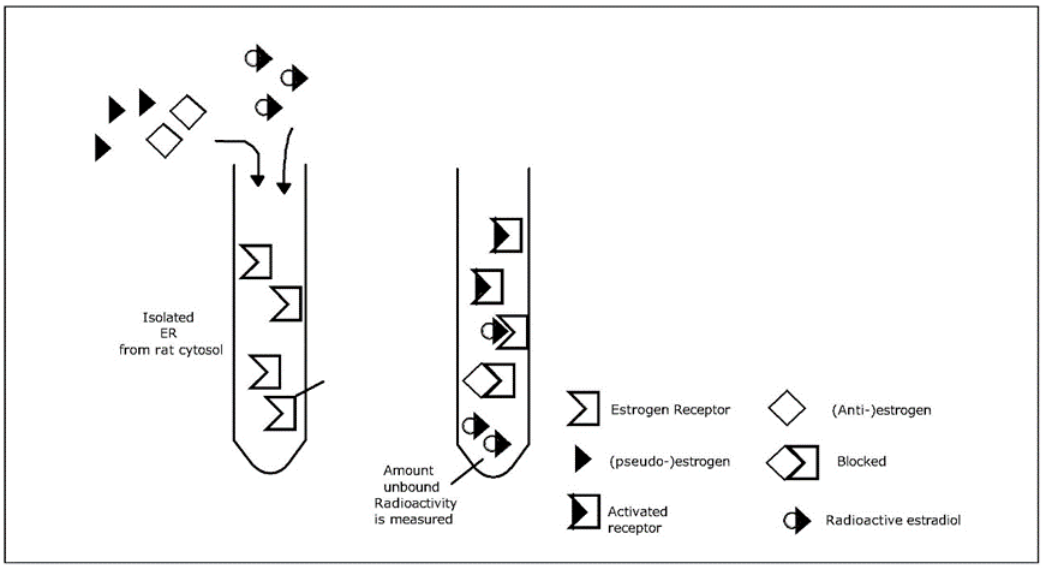


Figure 2. Principle of a radioactive ligand binding assay to determine binding of (anti-)estrogenic compounds to the estrogen receptor (ER). The ER is incubated with radiolabeled estradiol in combination with different concentrations of the test compound. If the compound is capable of binding to the ER, it will displace estradiol from the receptor. After separation of the free and bound estradiol, the amount of unbound radioactivity is measured. Increasing test concentrations of (anti-)estrogenic ER-binders will cause an increase in unbound radioactivity (and consequently a decrease in bound radioactivity). Redrawn from Murk et al. (2002) by Wilma Ijzerman.

Enzyme inhibition assays are used to determine if a test substance is capable to inhibit the enzymatic activity of a protein. Enzymatic activity is usually determined as the conversion rate of a substrate into a product. Enzyme inhibition is determined as a decrease in conversion rate, corresponding to lower concentrations of product and higher concentrations of substrate after different periods of incubation. Quantitative measures of substrate disappearance or product formation can be done by chemical analysis of the substrate or the product. Preferably, however, the reaction rate is measured by spectrophotometry or by fluorescence. This is achieved by performing the reaction with a substrate that has a specific colour or fluorescence by itself or that yields a product with a specific colour or fluorescence, in some cases after reaction with an additional indicator compound. A well-known example of an enzyme inhibition assay is the acetylcholinesterase inhibition assay (see section on Diagnosis - In vitro bioassays).

Cell cultures

Cell-based bioassays make use of cell cultures that are maintained in the laboratory. Cell culturing starts with mechanical or enzymatic isolation of single cells from a tissue (obtained from an animal or a plant). Subsequently, the cells are grown in cell culture medium, i.e. a liquid that contains all essential nutrients required for optimal cell growth (e.g. growth factors, vitamins, amino acids) and regulates the physicochemical environment of the cells (e.g. pH buffer, salinity). Typically, several types of cell cultures can be distinguished (Figure 3).

Primary cell cultures consist of cells that are directly isolated from a donor organism and are maintained in vitro. Typically, such cell cultures consist of either a cell suspension of non-adherent cells or a monolayer of adherent cells attached to a substrate (i.e. often the bottom of the culture vessel). The cells may undergo several cell divisions until the cell suspension becomes too dense or

the adherent cells grow on top of each other. The cells can then be further subcultured by transferring part of the cells from the primary culture to a new culture vessel containing fresh medium. This progeny of the primary cell culture is called a **cell line**, whereas the event of subculturing is called a **passage**. Typically, cell lines derived from primary cells undergo senescence and stop proliferating after a limited number (20-60) of cell divisions. Consequently, such a **finite cell line** can undergo only a limited number of passages. Primary cell cultures and their subsequent finite cell lines have the advantage that they closely resemble the physiology of the cells in vivo. The disadvantage of such cell cultures for toxicity testing is that they divide relatively slowly, require specific cell culturing conditions, and are finite. New cultures can only be obtained from new donor organisms, which is time-consuming, expensive, and may introduce genetic variation.

Alternatively, **continuous cell lines** have been established, which have an indefinite life span because the cells are immortal. Due to genetic mutations cells from a continuous cell line can undergo an indefinite number of cell divisions and behave like cancer cells. The immortalizing mutations may have been present in the original primary cell culture, if these cells were isolated from a malign cancer tumour tissue. Alternatively, the original finite cell line may have been transformed into a continuous cell line by introducing a viral or chemical induced mutation. The advantage of continuous cell lines is that the cells proliferate quickly and are easy to culture and to manipulate (e.g. by genetic modification). The disadvantage is that continuous cell lines have a different genotype and phenotype than the original healthy cells in vivo (e.g. have lost enzymatic capacity) and behave like cancer cells (e.g. have lost their differentiating capacities and ability to form tight junctions).

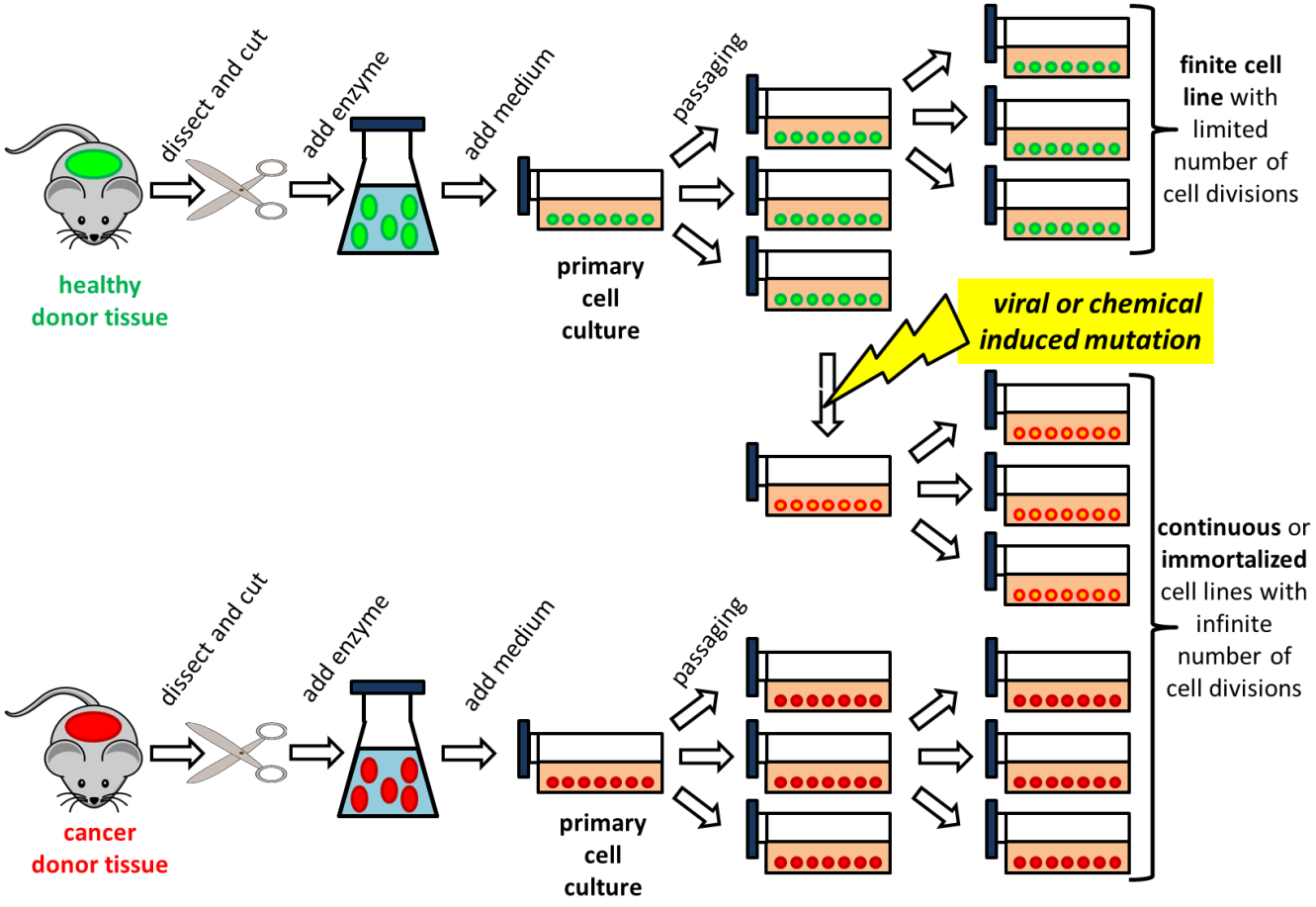


Figure 3: Different types of cell culturing, showing the establishment of a primary cell culture, a finite cell line, and a continuous cell line. See text for further explanation.

Differentiation models

To study the toxic effects of compounds in vitro, toxicologists prefer to use cell cultures that resemble differentiated, healthy cells rather than undifferentiated cancer cells. Therefore, **differentiation models** have gained increasing attention in in vitro toxicology in recent years. Such differentiation models are based on stem cells, which are cells that possess the potency to differentiate into

somatic cells. Stem cells can be obtained from embryonic tissues at different stages of normal development, each with their own potency to differentiate into somatic cells (see Figure 1 in Berdasco and Esteller, 2011). In the very early embryonic stage, cells from the "morula stage" (i.e. after a few cell divisions of the zygote) are **totipotent**, meaning that they can differentiate in all cell types of an organism. Later in development, cells from the inner cell mass of the trophoblast are **pluripotent**, meaning that they can differentiate in all cell types, except for extra-embryonic cells. During gastrulation, cells from the different germ layers (i.e. ectoderm, mesoderm, and endoderm) are **multipotent**, meaning that they can differentiate into a restricted number of cell types. Further differentiation results in precursor cells that are **unipotent**, meaning that they are committed to differentiate into a single ultimate differentiated cell type.

While remaining undifferentiated, in vitro embryonic stem cell (ESC) cultures can divide indefinitely, because they do not suffer from senescence. However, an ESC cell line cannot be considered as a continuous (or immortalized) cell line, because the cells contain no genetic mutations. ESCs can be differentiated into the cell type of interest by manipulating the cell culture conditions in such a way that specific signalling pathways are stimulated or inhibited in the same sequence as happens during in vivo cell type differentiation. Manipulation may consist of addition of growth factors, transcription factors, cytokines, hormones, stress factors, etc. This approach requires good understanding of which factors affect decision steps in the cell lineage of the cell type of interest.

Differentiation of ESCs into differentiated cells is not only applicable in in vitro toxicity testing, but also in drug discovery, regenerative medicine, and disease modelling. Still, the destruction of a human embryo for the purpose of isolation of - mainly pluripotent - human ESCs (hESCs) raises ethical issues. Therefore, alternative sources of hESCs have been explored. The isolation and subsequent in vitro differentiation of multipotent stem cells from amniotic fluid (collected during caesarean sections), umbilical cord blood, and adult bone marrow is a very topical field of research.

A revolutionary development in the field of non-embryonic stem cell differentiation models was the discovery that differentiated cells can be reprogrammed to undifferentiated cells with pluripotent capacities, called **induced pluripotent stem cells** (**iPSCs**) (Figure 4). In 2012, the Nobel Prize in Physiology or Medicine was awarded to John B. Gurdon and Shinya Yamanaka for this ground-breaking discovery. Reprogramming of differentiated cells isolated from an adult donor is obtained by exposing the cells to a mixture of reprogramming factors, consisting of transcription factors typical for pluripotent stem cells. The obtained iPSCs can be differentiated again (similar as ESCs) into any type of differentiated cells, for which the required conditions for cell lineage are known and can be simulated in vitro.

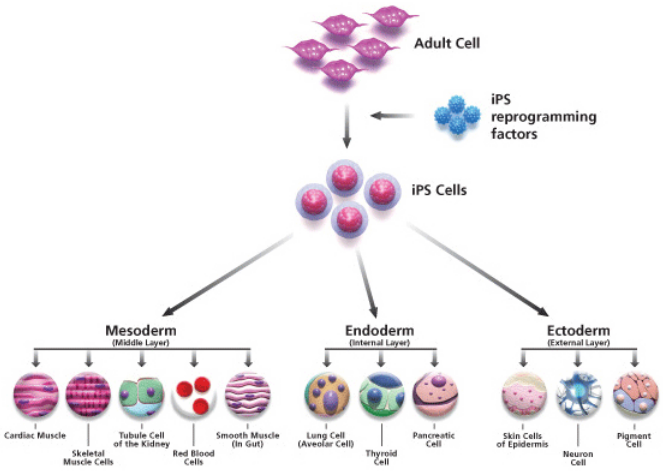


Figure 4: Principle of generating induced pluripotent stem cells (iPSCs) that can differentiate into any type of somatic cell. Source: https://beyondthedish.wordpress.com/2015/08/08/new-york-stem-cell-foundation-invents-robotic-platform-for-making-induced-pluripotent-stem-cells//

Whereas iPSC based differentiation models require a complete reprogramming of a differentiated somatic cell back to the stem cell level, **transdifferentiation** (or **lineage reprogramming**) is an alternative technique by which differentiated somatic cells can be transformed into another type of differentiated somatic cells, without undergoing an intermediate pluripotent stage. Especially fibroblast cell lines are known for their capacity to be transdifferentiated into different cell types, like neurons or adipocytes (Figure 5).

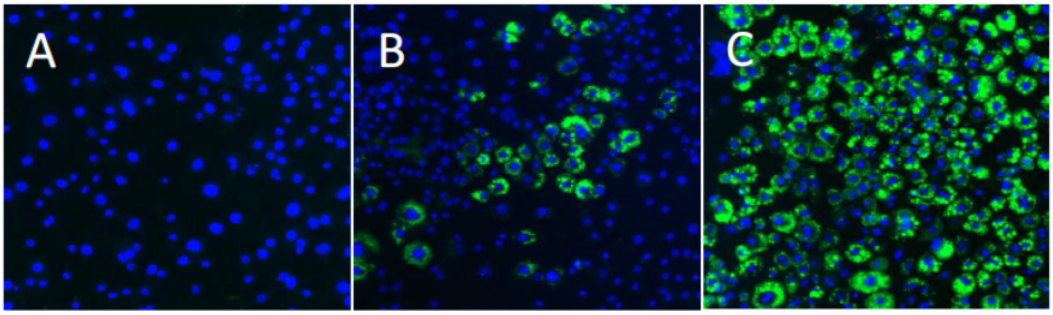


Figure 5. In vitro trans-differentiation of fibroblast cells from the 3T3-L1 cell line into mature adipocytes containing lipid vesicles (green). Each individual cell is visualized by nuclear staining (blue). A: undifferentiated control cells, B:cells exposed to an adipogenic cocktail consisting 3-isobutyl-1-methylxanthine, dexamethasone and insulin (MDI), C: cells exposed to MDI in combination with the PPAR gamma agonist troglitazone, an antidiabetic drug. Source: Vrije Universiteit Amsterdam-Dept. Environment & Health.

Cell-based bioassays

In cell-based in vitro bioassays, the cell cultures are exposed to test compounds or samples and their response is measured. In principle, all types of cell culture models discussed above can be used for in vitro toxicity testing. For reasons of time, money, and comfort, continuous cell lines are commonly used, but more and more often primary cell lines and iPSC-derived cell lines are used, for reasons of higher biological relevance. Endpoints that are measured in in vitro cell cultures exposed to toxic compounds typically range from effects on cell viability (measured as decreased mitochondrial functioning, increased membrane damage, or changes in cell metabolism; see section on Cytotoxicity) and cell growth to effects on cell kinetics (absorption, elimination and biotransformation of cell substrates), changes in the cell transcriptome, proteome or metabolome, or effects on cell-type dependent functioning. In addition, cell differentiation models can be used not only to study effects of compounds on differentiated cells, but also to study the effects on the process of cell differentiation per se by exposing the cells during differentiation.

A specific type of cell-based bioassays are the **reporter gene bioassays**, which are often used to screen individual compounds or complex mixtures extracted from environmental samples for their potency to activate or inactivate receptors that play a role in the expression of genes that play an important role in a specific path. Reporter gene bioassays make use of genetically modified cell lines or bacteria that contain an incorporated gene construct encoding for an easily measurable protein (i.e. the reporter protein). This gene construct is developed in such a way that its expression is triggered by a specific interaction between the toxic compound and a cellular receptor. If the receptor is activated by the toxic compound, transcription and translation of the reporter protein takes place, which can be easily measured as a change in colour, fluorescence, or luminescence (see section on Diagnosis - In vitro bioassays).

Future developments

Although there is a societal need for a non-toxic environment, there is also a societal demand to **R**educe, **R**efine and **R**eplace animal studies (three R principles). Replacement of animal studies by in vitro tests requires that the obtained in vitro results are indicative and predictive for what happens in the in vivo situation. It is obvious that a cell culture consisting of a single cell type is not comparable to a complex organism. For instance, toxicokinetic aspects are hardly taken into account in cell-based bioassays. Although some cells might have metabolic capacities, processes like adsorption, distribution, and elimination are not represented as exposure is usually directly on the cells. Moreover, cell cultures often lack repair mechanisms, feedback loops, and any other interaction with other cell types/tissues/organs as found in intact organisms. To expand the scope of in vitro - in vivo extrapolation (IVIVE), more complex in vitro models are developed nowadays that have a closer resemblance to the in vivo situation. For instance, whereas cell culturing was traditionally done in 2D monolayers (i.e. in layers of 1 cell thickness), 3D cell culturing is gaining ground. The advantage of 3D culturing is that it represents a more realistic type of cell growth, including cell-cell interactions, polarization, differentiation, extracellular matrix, diffusion gradients, etc. For epithelial cells (e.g. lung cells), such 3D cultures can even be grown at the air-liquid interphase reflecting the in vivo situation. Another development is cell co-culturing where different cell types are cultured together in a cell culture. For instance, two cell types that interact in an organ can be co-cultured. Alternatively, a differentiated cell type that has poor metabolic capacity can be co-cultured with a liver cell in order to take possible detoxification or bioactivation after biotransformation into account. The latest development in increasing complexity

in in vitro test systems are so-called organ-on-a-chip devices, in which different cell types are co-cultured in miniaturized small channels. The cells can be exposed to different flows representing for instance the blood stream, which may contain toxic compounds (see for instance video clips at https://wyss.harvard.edu/technology/human-organs-on-chips/). Based on similar techniques, even human body-on-a-chip devices can be constructed. Such chips contain different miniaturized compartments containing cell co-cultures representing different organs, which are all interconnected by different channels representing a microfluid circulatory system (Figure 6). Although such devices are in their infancies and regularly run into impracticalities, it is to be expected that these innovative developments will play their part in the near future of toxicity testing.

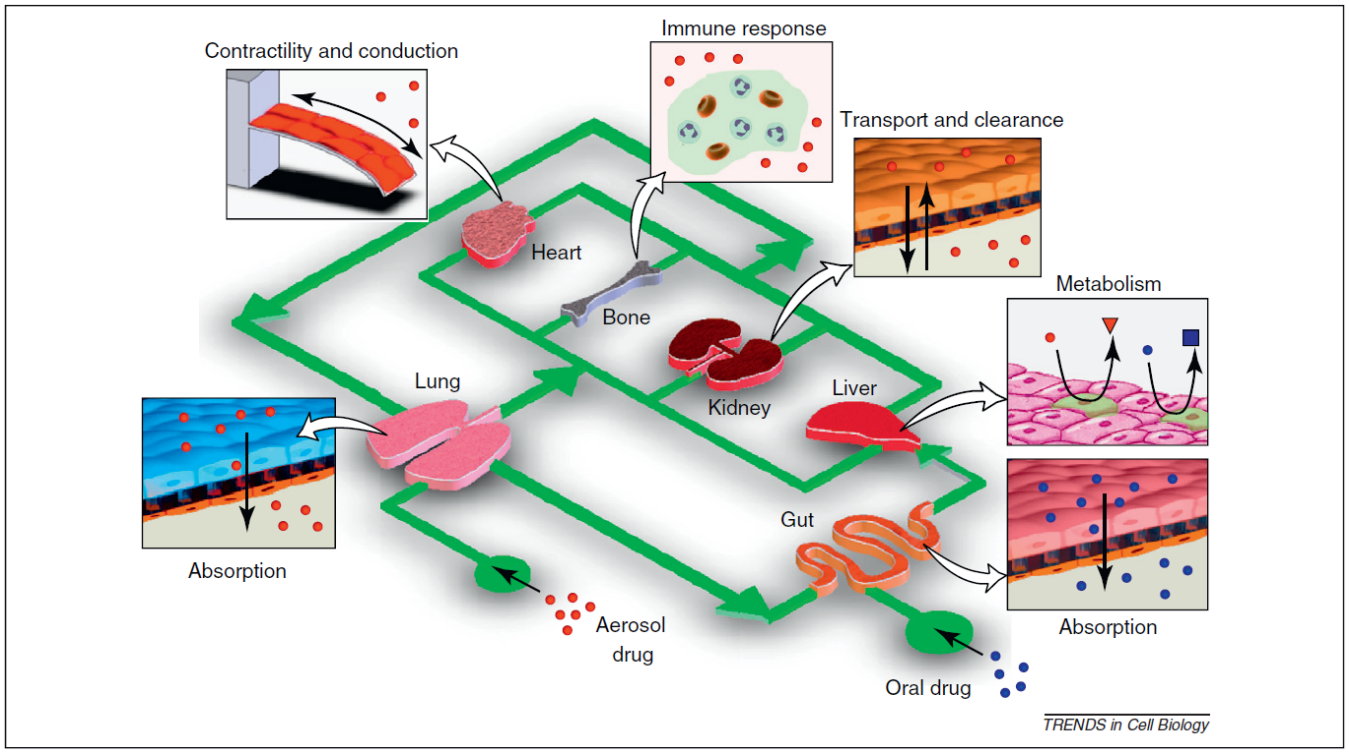


Figure 6: The human-on-a-chip device, showing miniaturized compartments (or biomimetic microsystems) containing (co-)cultures representing different organs, interconnected by a microfluidic circulatory system. Compartments are connected in a physiologically relevant manner to reflect complex, dynamic ADME processes and to allow toxicity evaluation. In this example, an integrated system of microengineered organ mimics (lung, heart, gut, liver, kidney and bone) is used to study the absorption of inhaled aerosol substances (red) from the lung to microcirculation, in relation to their cardiotoxicity (e.g. changes in heart contractility or conduction), transport and clearance in the kidney, metabolism in the liver, and immune-cell contributions to these responses. To investigate the effects of oral administration, substances can also be introduced into the gut compartment (blue).

4.3.8. Question 1

What is the principle of a ligand binding assay?

4.3.8. Question 2

What is the principle of an enzyme inhibition assay?

4.3.8. Question 3

What is the principle of a reporter gene bioassay?

4.3.8. Question 4

What are the advantages and disadvantages of primary cell cultures versus continuous cell lines?

4.3.8. Question 5

What is the difference between embryonic stem cells and induced pluripotent stem cells?

4.3.9. Human toxicity testing - I. General aspects

In preparation

4.3.9. Human toxicity testing - II. In vitro tests

(Draft)

Author: Nelly Saenen

Reviewers: Karen Smeets, Frank Van Belleghem

Learning objectives:

You should be able to

- argue the need for alternative test methods for toxicity
- list commonly used *in vitro* cytotoxicity assays and explain how they work
- describe different types of toxicity to skin and *in vitro* test methods to assess this type of toxicity

Keywords: *In vitro*, toxicity, cytotoxicity, skin

Introduction

Toxicity tests are required to assess potential hazards of new compounds to humans. These tests reveal species-, organ- and dose-specific toxic effects of the compound under investigation. Toxicity can be observed by either *in vitro* studies using cells/cell lines (see section on in vitro bioassay) or by *in vivo* exposure on laboratory animals; and involves different durations of exposure (acute, subchronic, and chronic). In line with Directive 2010/63/EC on the protection of animals used for scientific purposes, it is encouraged to use alternatives to animal testing (OECD: alternative methods for toxicity testing). The first step towards replacing animals is to use *in vitro* methods which can be used to predict acute toxicity. In this chapter, we present acute *in vitro* cytotoxicity tests (= quality of being toxic to cells) and skin corrosive, irritant, phototoxic, and sensitivity tests as skin is the largest organ of the body.

1. *Cytotoxicity tests*

The cytotoxicity test is one of the biological evaluation and screening tests *in vitro* to observe cell viability. Viability levels of cells are good indicators of cell health. Conventionally used tests for cytotoxicity include dye exclusion or uptake assays such as Trypan Blue Exclusion (TBE) and Neutral Red Uptake (NRU).

The **TBE test** is used to determine the number of viable cells present in a cell suspension. Live cells possess intact cell membranes that exclude certain dyes, such as trypan blue, whereas dead cells do not. In this assay, a cell suspension incubated with serial dilutions of a test compound under study is mixed with the dye and then visually examined. A viable cell will have a clear cytoplasm whereas a nonviable cell will have a blue cytoplasm. The number of viable and/or dead cells per unit volume is determined by light microscopy using a hemacytometer counting chamber (Figure 1). This method is simple, inexpensive and a good indicator of membrane integrity but counting errors (~10%) can occur due to poor dispersion of cells, or improper filling of counting chamber.

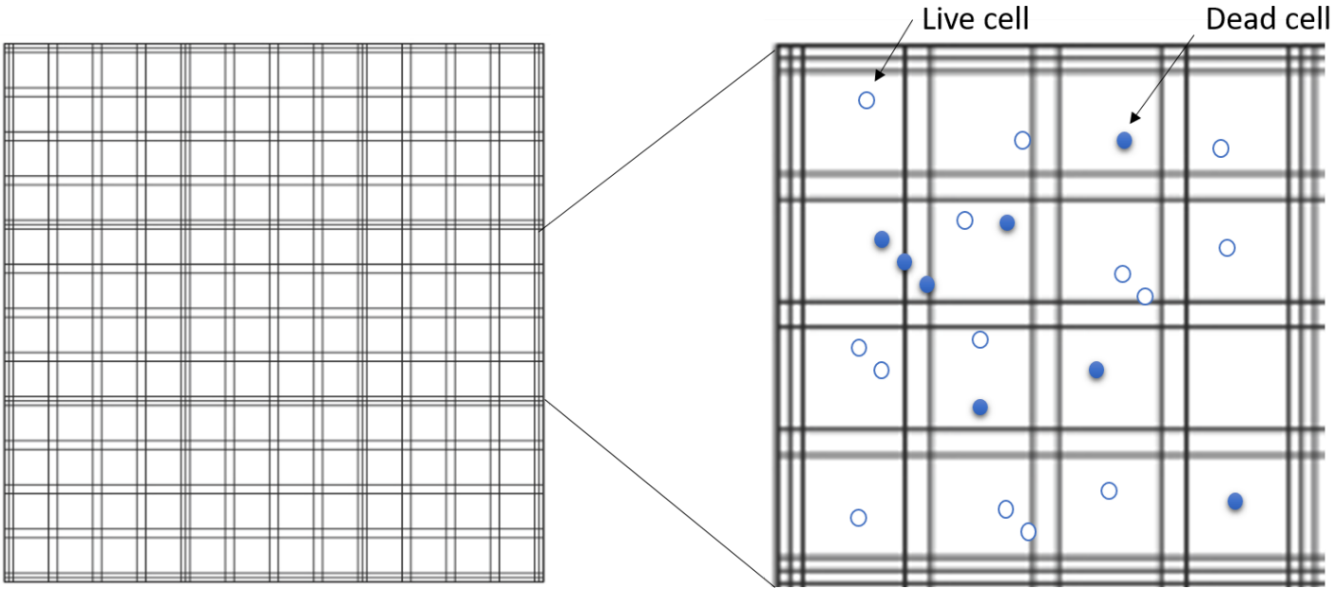


Figure 1: Graphical view hemacytometer counting chamber and illustration of viable and non-viable cells.

The **NRU assay** is an approach that assesses the cellular uptake of a dye (Neutral Red) in the presence of a particular substance under study (see e.g.Figure 1 in Repetto et al., 2008). This test is based on the ability of viable cells to incorporate and bind neutral red in the lysosomes, a process based on universal structures and functions of cells (e.g. cell membrane integrity, energy production and metabolism, transportation of molecules, secretion of molecules). Viable cells can take up neutral red via active transport and incorporate the dye into their lysosomes while non-viable cells cannot. After washing, viable cells can release the incorporated dye under acidified-extracted conditions. The amount of released dye can be measured by the use of spectrophotometry.

Nowadays, colorimetric assays to assess cell viability have become popular. For example, the **MTT assay** (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) tests cell viability by assessing activity of mitochondrial enzymes. NAD(P)H-dependent oxidoreductase enzymes, which under defined conditions reflect the number of viable cells, are capable of reducing the yellow formazan salt dye into an insoluble purple crystalline product. After solubilizing the end product using dimethyl sulfoxide (DMSO), the product can be quantitatively measured by light absorbance at a specific wavelength. This method is easy to use, safe and highly reproducible. One disadvantage is that MTT formazan is insoluble so DMSO is required to solubilize the crystals.

2. Skin corrosion and irritation

Skin corrosion refers to the production of irreversible damage to the skin; namely, visible necrosis (= localized death of living cells, see section on cell death) through the epidermis and into the dermis occurring after exposure to a substance or mixture. Skin irritation is a less severe effect in which a local inflammatory reaction is observed onto the skin after exposure to a substance or mixture. Examples of these substances are detergents and alkalis which commonly affect hands.

The identification and classification of irritant substances has conventionally been achieved by means of skin or eye observation *in vivo*. Traditional animal testing used rabbits because of its thin skin. In the Draize test, for example, the test substance is applied to the eye or shaved skin of a rabbit, and covered for 24h. After 24 and 72h, the eye or skin is visually examined and graded subjectively based on appearance of erythema and edema. As these *in vivo* tests have been heavily criticized, they are now being phased out in favor of *in vitro* alternatives.

The Skin Corrosion Test (SCT) and Skin Irritation Test (SIT) are *in vitro* assays that can be used to identify whether a chemical has the potential to corrode or irritate skin. The method uses a three-dimensional (3D) human skin model (Episkin model) which comprises of a main basal, supra basal, spinous and granular layers and a functional stratum corneum (the outer barrier layer of skin). It involves topical application of a test substance and subsequent assessment of cell viability (MTT assay). Test compounds considered corrosive or irritant are identified by their ability to decrease cell viability below the defined threshold level (lethal dose by which 50% of cells are still viable - LD_{50}).

3. Skin phototoxicity

Phototoxicity (photoirritation) is defined as a toxic response that is elicited after the initial exposure of skin to certain chemicals and subsequent exposure to light (e.g. chemicals that absorb visible or ultraviolet (UV) light energy that induces toxic molecular changes).

The 3T3 NRU PT assay is based on an immortalised mouse fibroblast cell line called Balb/c 3T3. It compares cytotoxicity of a chemical in the presence or absence of a non-cytotoxic dose of simulated solar light. The test expresses the concentration-dependent reduction of the uptake of the vital dye neutral red when measured 24 hours after treatment with the chemical and light irradiation. The exposure to irradiation may alter cell surface and thus may result in decreased uptake and binding of neutral red. These differences can be measured with a spectrophotometer.

4. Skin sensitisation

Skin sensitisation is the regulatory endpoint aiming at the identification of chemicals able to elicit an allergic response in susceptible individuals. In the past, skin sensitisation has been detected by means of guinea pigs (e.g. guinea pig maximisation test and the Buehler occluded patch tests) or murine (e.g. murine local lymph node assay). The latter is based upon quantification of T-cell proliferation in the draining lymph nodes behind the ear (auricular) of mice after repeated topical application of the test compound.

The key biological events (Figure 2) underpinning the skin sensitisation process are well established and include:

1. haptenation, the covalent binding of the chemical compounds (haptens) to skin proteins (key event 1);
2. signaling, the release of pro-inflammatory cytokines and the induction of cyto-protective pathways in keratinocytes (key event 2);
3. the maturation, and mobilisation of dendritic cells, immuno-competent cells in the skin (key event 3);
4. migration of dendritic cells, movement of dendritic cells bearing hapten-protein comples from skin to draining local lymph node;
5. the antigen presentation to naïve T-cells and proliferation (clonal expansion) of hapten-peptide specific T-cells (key event 4)

Figure in preparation

Figure 2: Key biological events in skin sensitisation. Figure adapted from …

Today a number of non-animal methods addressing each a specific key mechanism of the induction phase of skin sensitisation can be employed. These include the Direct Peptide Reactivity Assay (DPRA), ARE-Nrf2 Luciferase Test Method: KeratinoSens, Human Cell Line Activation Test (h-CLAT), U937 cell line activation test (U-SENS), and Interleukin-8 Reporter Gene assay (IL-8 Luc assay). Detailed information of these methods can be found on OECD site: skin sensitization.

References

EUR-lex site. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32010L0063

OECD site. Alternative methods for toxicity testing. https://ec.europa.eu/jrc/en/eurl/ecvam/alternative-methods-toxicity-testing

Episkin model. http://www.episkin.com/Episkin

OECD site. Skin sensitization. https://ec.europa.eu/jrc/en/eurl/ecvam/alternative-methods-toxicity-testing/validated-test-methods/skin-sensitisation

Repetto, G., Del Peso, A., Zurita, J.L. (2008). Neutral red uptake assay for the estimation of cell viability/cytotoxicity.Nature protocols *3*(7), 1125.

4.3.9. Human toxicity testing - III. Carcinogenicity assays

(Draft)

Author: Jan-Pieter Ploem

Reviewers: Frank van Belleghem

Learning objectives:

You should be able to

- explain the different approaches used for carcinogen testing.
- list some advantages and disadvantages of the different methods
- understand the difference between GTX and NGTX compounds, and its consequence regarding toxicity testing

Introduction

The term "carcinogenicity" refers to the property of a substance to induce or increase the incidence of cancer after inhalation, ingestion, injection or dermal application.

Traditionally, carcinogens have been classified according to their mode of action (MoA). Compounds directly interacting with DNA, resulting in DNA-damage or chromosomal aberrations are classified as genotoxic (GTX) carcinogens. Non-genotoxic (NGTX) compounds do not directly affect DNA and are believed to affect gene expression, signal transduction, disrupt cellular structures and/or alter cell cycle regulation.

The difference in mechanism of action between GTX and NGTX require a different approach in many cases.

Genotoxic carcinogens

Genotoxicity itself is considered to be an endpoint in its own right. The occurrence of DNA-damage can be observed/determined quit easily by a variety of methods based on both bacterial and mammalian cells. Often a tiered testing method is used to evaluate both heritable germ cell line damage and carcinogenicity.

Currently eight *in vitro* assays have been granted OECD guidelines, four of which are commonly used.

- The Ames test

The gold standard for genotoxicity testing is the Ames test, a test that has been developed in the early seventies. The test evaluates the potential of a chemical to induce mutations (base pair substitutions, frame shift induction, oxidative stress, etc.) in *Salmonella typhimurium*. During the safety assessment process, it is the first test performed unless deemed unsuitable for specific reasons (e.g. during testing for antibacterial substances). With a sensitivity of 70-90% it is a relatively good predictor of genotoxicity.

The principle of the test is fairly simple. A bacterial strain, with a genetic defect, is placed on minimal medium containing the chemical in question. If mutations are induced, the genetic defect in some cells will be restored, thus rendering the cells able to synthesize the deficient amino acid caused by the defect in the original cells.

- *Escherichia coli* reverse mutation assay

The Ames assay is basically a bacterial reverse mutation assay. In this case different strains of *E. coli*, which are deficient in both DNA-repair and an amino acid, are used to identify genotoxic chemicals. Often a combination of different bacterial strains is used to increase the sensitivity as much as possible.

- *In vitro* mammalian chromosome aberration assay

Chromosomal mutations can occur in both somatic cells and in germ cells, leading to neoplasia or birth and developmental abnormalities respectively. There are two types of chromosomal mutations:

Structural changes: stable aberrations such as translocations and inversions, and unstable aberrations such as gaps and breaks.

Numerical changes: aneuploidy (loss or gain of chromosomes) and polyploidy (multiples of the diploid chromosome complement).

To perform the assay, mammalian cells are exposed *in vitro* to the potential carcinogen and then harvested. Through microscopy the frequency of aberrations is determined. The chromosome aberration can be and is performed with both rodent and human cells which is an interesting feat regarding the translational power of the assay.

- *In vitro* mammalian cell gene mutation test

This mammalian genotoxicity assay utilizes the HPRT gene, a X-chromosome located reporter gene. The test relies on the fact that cells with an intact HPRT gene are susceptible to the toxic effects of 6-thioguanine, while mutants are resistant to this purine analogue. Wild-type cells are sensitive to the cytostatic effect of the compound while mutants will be able to proliferate in the presence of 6-thioguanine.

- Micronucleus test

Next to the four mentioned assays, there is a more recent developed test that already has proven to be a valuable resource for genotoxicity testing. The test provides an alternative to the chromosome aberration assay but can be evaluated faster. The test allows for automated measurement as the analysis of the damage proves to be less subjective. Micronuclei are "secondary" nuclei formed as a result of aneugenic or clastogenic damage.

It is important to note that these assays are all described from an *in vivo* perspective. However, an *in vivo* approach can always be used, see two-year rodent assay. In that case, live animals are exposed to the compound after which specific cells are harvested. The advantage of this approach is the presence of the natural niche in which susceptible cells normally grow, resulting in a more relevant range of effects. The downside of *in vivo* assays is the current ethical pressure on these kind of methods. Several instances actively promote the development and usage of *in vitro* or simply non-animal alternative methods.

- Two-year rodent carcinogenicity assay

For over 50 years, 2-year rodent carcinogenicity assay has been the golden standard for carcinogenicity testing. The assay relies on the exposure to a compound during a major part of an organism's lifespan. During the further development of the assay, a 2-species/2-gender setup became the preferred method, as some compounds showed different results in e.g. rats and mice and even between male and female individuals.

For this approach model organisms are exposed for two years to a compound. Depending on the possible mode of exposure (i.e. inhalation, ingestion, skin/eye/... contact) when the compounds enters the relevant industry, a specific mode of exposure towards the model is chosen. During this time period the health of the model organism is documented through different parameters. Based on this a conclusion regarding the compound is drawn.

Non-genotoxic carcinogens

Carcinogens not causing direct DNA-damage are classified as NGTX compounds. Due to the fact that there are a large number of potential malign pathways or effects that could be induced, the identification of NGTX carcinogens is significantly more difficult compared to GTX compounds.

The two-year rodent carcinogenicity assay is one of the assays capable of accurately identify NGTX compounds. The use of transgenic models has greatly increased the sensitivity and specificity of this assay towards both groups of carcinogens while also improving the refinement of the assay by shortening the required time to formulate a conclusion regarding the compound.

An *in vitro* method to identify NGTX compounds is rare. Not many alternative assays are able to cope with the vast variety of possible effects caused by the compounds resulting in many false negatives. However, cell morphology based methods such as the cell transformation assay, can be a good start in developing methods for this type of carcinogens.

References

Stanley, L. (2014). Molecular and Cellular Toxicology. p. 434.

4.3.10. Environmental epidemiology - I. Basic principles and study designs

Authors: Eva Sugeng and Lily Fredrix

Reviewers: Ľubica Murínová and Raymond Niesink

Learning objectives:

You should be able to

- describe and apply definitions of epidemiologic research.
- name and identify study designs in epidemiology, describe the design, and advantages and disadvantages of the design.

1. Definitions of epidemiology

Epidemiology (originating from Ancient Greek: *Epi* -upon, *demos* - people, *logos* - the study of) is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the prevention and control of health problems (Last, 2001). Epidemiologists study human populations with measurements at one or more points in

time. When a group of people is followed over time, we call this a **cohort** (originating from Latin: *cohors* (Latin), a group of Roman soldiers). In epidemiology, the relationship between a **determinant** or **risk factor** and health **outcome** - variable is investigated. The **outcome** variable mostly concerns **morbidity**: a disease, e.g. lung cancer, or a health parameter, e.g. blood pressure, or **mortality**: death. The **determinant** is defined as *a collective or individual **risk factor** (or set of factors) that is (causally) related to a health condition, outcome, or other defined characteristic. In human health - and, specifically, in diseases of complex etiology - sets of determinants often act jointly in relatively complex and long-term processes* (International Epidemiological Association, 2014).

The people that are subject of interest are the **target population**. In most cases, it is impossible and unnecessary to include all people from the target population and therefore, a **sample** will be taken from the target population, which is called the **study population**. The **sample** is ideally representative of the target population (Figure 1). To get a representative sample, it is possible to recruit subjects at random.

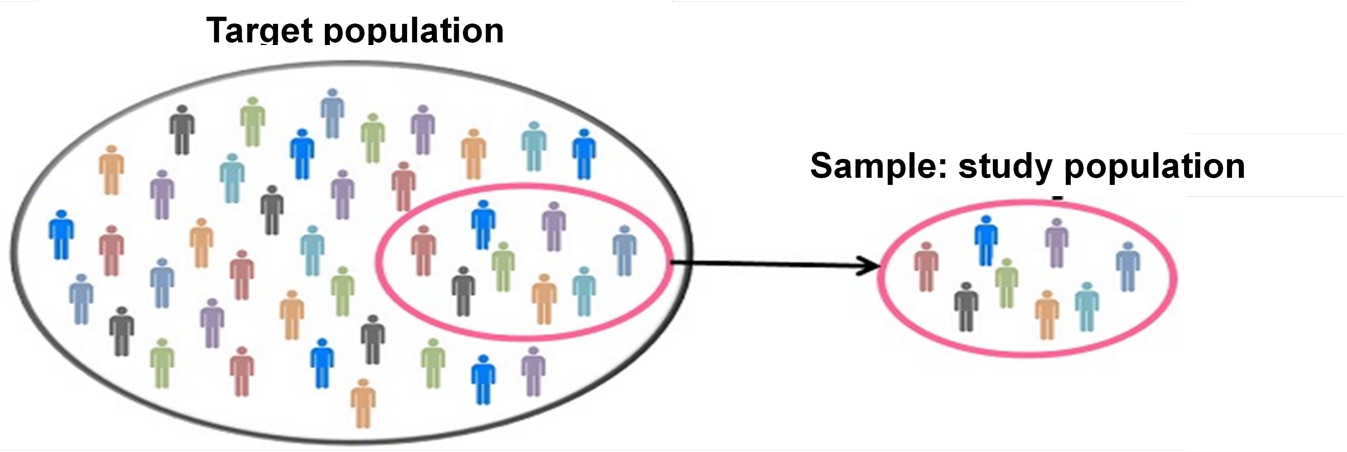


Figure 1: On the left, the target population is presented and from this population, a representative sample is drawn, including all types of individuals from the target population.

2. Study designs

Epidemiologic research can either be **observational** or **experimental** (Figure 2). Observational studies do not include interference (e.g. allocation of subjects into exposed / non-exposed groups), while experimental studies do. With regard to observational studies **analytical** and **descriptive** studies can be distinguished. Descriptive studies describe the determinant(s) and outcome without making comparisons, while analytical studies compare certain groups and derive inferences.

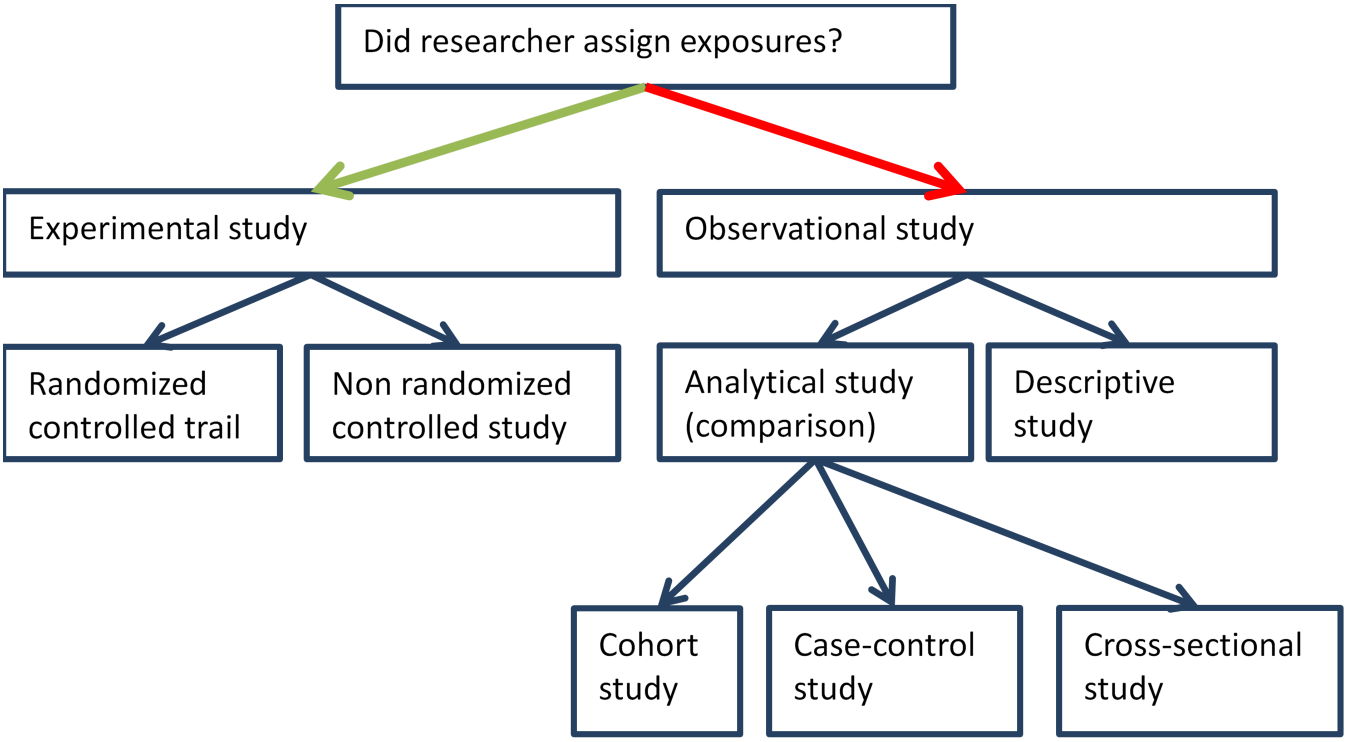


Figure 2: The types of study designs, the branch on the left includes an exposure/intervention assigned by the researcher, while the branch on the right is observational, and does not include an exposure/intervention assigned by the researcher.

2.1 Observational studies

2.1.1. Cross-sectional study

In a **cross-sectional study**, **determinant** and **outcome** are measured at the same time. For example, pesticide levels in urine (**determinant**) and hormone levels in serum (**outcome**) are collected at one point in time. The design is quick and cheap because all measurements take place at the same time. The drawback is that the design does not allow to conclude about **causality**, that is whether the **determinant** precedes the **outcome**, it might be the other way around or caused by another factor (lacking **Hill's criterion for causality** temporality, Box 1). This study design is therefore mostly hypothesis generating.

2.1.2 Case-control study

In a **case-control study**, the sample is selected based on the **outcome**, while the **determinant** is measured in the past. In contrast to a cross-sectional study, this design can include measurements at several time points, hence it is a longitudinal study. First, people with the disease (**cases**) are recruited and then **matched controls** (people not affected by the disease), comparable with regard to e.g. age, gender and geographical region, are involved into the study. Important is that controls have the same risk to develop the disease as the cases. The determinant is collected **retrospectively**, meaning that participants are asked about exposure in the past.

The retrospective character of the design poses a risk for **recall bias** when people are asked about events that happened in the past, they might not remember them correctly. **Recall bias** is a form of **information bias**, when a measurement error results in misclassification. **Bias** is defined as a systematic deviation of results or inferences from the truth (International Epidemiological Association, 2014). One should be cautious to draw conclusions about **causality** with the case-control study design. According to **Hill's criterion temporality** (see Box 1), the exposure precedes the outcome, but because the exposure is collected retrospectively, the evidence may be too weak to draw conclusions about a **causal relationship**. The benefits are that the design is suitable for research on diseases with a **low incidence** (in a **prospective cohort study** it would result in a low number of cases), and for research on diseases with a long **latency period**, that is the time that exposure to the determinant can result in the disease (in a **prospective cohort study**, it would take many years to follow-up participants until the disease develops).

An example of a case-control study in environmental epidemiology

Hoffman et al. (2017) investigated papillary thyroid cancer (PTC) and exposure to flame retardant chemicals (FRs) in the indoor environment. FRs are chemicals which are added to household products in order to limit the spread of fire, but can leach to house dust where residents can be exposed to the contaminated house dust. FRs are associated with thyroid disease and thyroid cancer. In this case-control study, PTC cases and matched cases were recruited (**outcome**), and FR exposure (**determinant**) was assessed by measuring FRs in the house dust of the participants. The study showed that participants with higher exposure to FRs (bromodiphenyl ether-209 concentrations above the median level) had 2.3 more **odds** (see section Quantifying disease and associations) on having PTC, compared to participants with lower exposure to FRs (bromodiphenyl ether-209 concentrations below the median level).

2.1.3 Cohort study

A **cohort study**, another type of a longitudinal study, includes a group of individuals that are followed over time in the future (**prospective**) or that will be asked about the past (**retrospective**). In a prospective cohort study, the **determinant** is measured at the start of the study and the **incidence** of the disease is calculated after a certain time period, the **follow-up**. The study design needs to start with people who are **at risk** for the disease, but not yet affected by the disease. Therefore, the prospective study design allows to conclude that there may be a **causal relationship**, since the health outcome follows the determinant in time (**Hill's criterion temporality**). However, interference of other factors is still possible, see paragraph 3 about confounding and effect modification. It is possible to look at more than 1 health outcome, but the design is less suitable for diseases with **a low incidence** or with a long **latency period**, because then you either need a large study population to have enough cases, or need to follow the participants for a long time to measure cases. A major issue with this study design is **attrition** (loss to follow-up), it means to what extent do participants drop out during the study course. **Selection bias** can occur when a certain type of participants drops out more often, and the research is conducted with a selection of the target population. **Selection bias** can also occur at the start of a study, when some members of the target population are less likely to be included in the study population in comparison to other members and the sample therefore is not representative of the target population.

An example of a prospective cohort study

De Cock et al. (2016) present a prospective cohort study investigating early life exposure to chemicals and health effects in later life, the LInking EDCs in maternal Nutrition to Child health (LINC study). For this, over 300 pregnant women were recruited during pregnancy. Prenatal exposure to chemicals was measured in, amongst others, cord blood and breast milk and the children were followed over time, measuring, amongst others, height and weight status. For example, prenatal exposure to dichlorodiphenyl-dichloroethylene (DDE), a metabolite of the pesticide dichlorodiphenyl-trichloroethane (DDT), was assessed by measuring DDE in umbilical cord blood, collected at delivery. During the first year, the body mass index (BMI), based on weight and height, was monitored. DDE levels in umbilical cord blood were divided into 4 equal groups, called quartiles. Boys with the lowest DDE concentrations (the first quartile) had a higher BMI growth curve in the first year, compared to boys with the highest concentrations DDE (the fourth quartile) (De Cock et al., 2016).

2.1.4 Nested case-control study

When a **case-control study** is carried out within a **cohort study**, it is called a **nested case-control study**. Cases in a cohort study are selected, and **matching** non-cases are selected as controls. This type of study design is useful in case of a low amount of cases in a prospective cohort study.

An example of a nested case-control study

Engel et al. (2018) investigated attention-deficit hyperactivity disorder (ADHD) in children in relation to prenatal phthalate exposure. Phthalates are added to various consumer products to soften plastics. Exposure occurs during ingestion, inhalation or dermal absorption and sources are for example plastic packaging of food, volatile household products and personal care products (Benjamin et al., 2017). Engel et al. (2018) carried out a nested case-control study within the Norwegian Mother and Child Cohort (MoBa). The cohort included 112,762 mother-child pairs of which only a small amount of cases with a clinical ADHD diagnosis. A total of 297 cases were randomly sampled from registrations of clinically ADHD diagnoses. In addition, 553 controls without ADHD were randomly sampled from the cohort. Phthalate metabolites were measured in maternal urine collected at midpregnancy

and concentrations were divided into 5 equal groups, called quintiles. Children of mothers in the highest quintile of the sum of metabolites of the phthalate bis(2-ethylhexyl) phthalate (DEHP) had 2.99 (95%CI: 1.47-5.49) more **odds** (see chapter Quantifying disease and associations) of an ADHD diagnosis in comparison to the lowest quintile.

2.2 Experimental studies

2.2.1 (Non) Randomized controlled trials

A **randomized controlled trial** (RCT) is an experimental study in which participants are randomly assigned to an intervention group or a control group. The intervention group receives an intervention or treatment, the control group receives nothing, usual care or a placebo. **Clinical trials** that test the effectiveness of medication are an example of an RCT. If the assignment of participants to groups is not randomized, the design is called a non-randomized controlled trial. The latter design provides less strength of evidence.

When groups of people instead of individuals, are randomized, the study design is called a **cluster-randomized controlled trial**. This is, for example, the case when classrooms with children at school are randomly assigned to the intervention- and control group. Variations are used to switch groups between the intervention and control group. For example, a **crossover design** makes it possible that people are both intervention group and control group in different phases of the study. In order to not restrain the benefits of the intervention to the control group, a **waiting list design** makes the intervention available to the control group after the research period.

An example of an experimental study

An example of an experimental study design within environmental research is the study of Bae and Hong (2015). In a randomized crossover trial, participants had to drink beverages either from a BPA containing can, or a BPA-free glass bottle. Besides BPA levels in urine, blood pressure was measured after exposure. The crossover design included 3 periods, with either drinking only canned beverages, both canned and glass-bottled beverages or only glass-bottled beverages. BPA concentration was increased with 1600% after drinking canned beverages in comparison to drinking from glass bottles.

3. Confounding and effect modification

Confounding occurs when a third factor influences both the **outcome** and the **determinant** (see Figure 3). For example, the number of cigarettes smoked is positively associated with the prevalence of esophageal cancer. However, the number of cigarettes smoked is also positively associated with the amount of standard glasses alcohol consumption. Besides, alcohol consumption is a **risk factor** for esophageal cancer. Alcohol consumption is therefore a **confounder** in the relationship smoking and esophageal cancer. One can correct for confounders in the statistical analysis, e.g. using stratification (results are presented for the different groups separately).

Effect modification occurs when the association between exposure/determinant and outcome is different for certain groups (Figure 3). For example, the risk of lung cancer due to asbestos exposure is about ten times higher for smokers than for non-smokers. A solution to deal with effect modification is **stratification as well**.

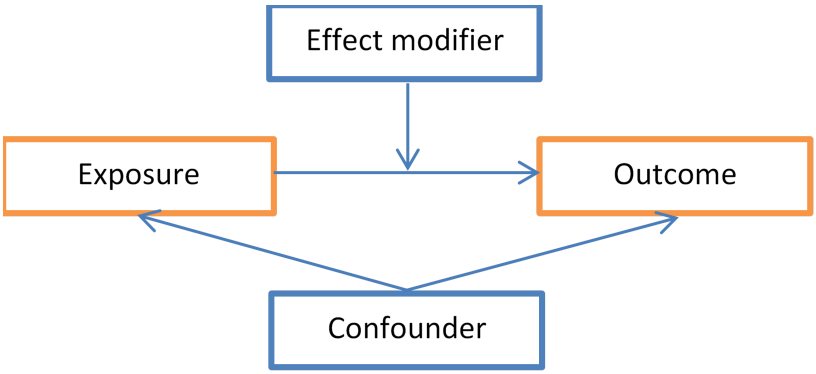


Figure 3: Confounding and effect modification in an association between exposure and **outcome**. A confounder has associations

with both the exposure/determinant and the outcome. An effect modifier alters the association between the exposure/determinant and the outcome.

Box 1: Hill's criteria for causation

With epidemiological studies it is often not possible to determine a causal relationship. That is why epidemiological studies often employ a set of criteria, the Hill's criteria of causation, according to Sir Austin Bradford Hill, that need to be considered before conclusions about causality are justified (Hill, 1965).

1. Strength: stronger associations are more reason for causation.
2. Consistency: causation is likely when observations from different persons, in different populations and circumstances are consistent.
3. Specificity: specificity of the association is reason for causation.
4. Temporality: for causation the determinant must precede the disease.
5. Biological gradient: is there biological gradient between the determinant and the disease, for example, a dose-response curve?
6. Plausibility: is it biological plausible that the determinant causes the disease?
7. Coherence: coherence between findings from laboratory analysis and epidemiology.
8. Experiment: certain changes in the determinant, as if it was an experimental intervention, might provide evidence for causal relationships.
9. Analogy: consider previous results from similar associations.

References

Bae, S., Hong, Y.C. (2015). Exposure to bisphenol a from drinking canned beverages increases blood pressure: Randomized crossover trial. Hypertension 65, 313-319. https://doi.org/10.1161/HYPERTENSIONAHA.114.04261

Benjamin, S., Masai, E., Kamimura, N., Takahashi, K., Anderson, R.C., Faisal, P.A. (2017). Phthalates impact human health: Epidemiological evidences and plausible mechanism of action. Journal of Hazardous Materials 340, 360-383. https://doi.org/10.1016/j.jhazmat.2017.06.036

De Cock, M., De Boer, M.R., Lamoree, M., Legler, J., Van De Bor, M. (2016). Prenatal exposure to endocrine disrupting chemicals and birth weight-A prospective cohort study. Journal of Environmental Science and Health - Part A Toxic/Hazardous Substances and Environmental Engineering 51, 178-185. https://doi.org/10.1080/10934529.2015.1087753

De Cock, M., Quaak, I., Sugeng, E.J., Legler, J., Van De Bor, M. (2016). Linking EDCs in maternal Nutrition to Child health (LINC study) - Protocol for prospective cohort to study early life exposure to environmental chemicals and child health. BMC Public Health 16: 147. https://doi.org/10.1186/s12889-016-2820-8

Engel, S.M., Villanger, G.D., Nethery, R.C., Thomsen, C., Sakhi, A.K., Drover, S.S.M., … Aase, H. (2018). Prenatal phthalates, maternal thyroid function, and risk of attention-deficit hyperactivity disorder in the Norwegian mother and child cohort. Environmental Health Perspectives. https://doi.org/10.1289/EHP2358

Hill, A.B. (1965). The Environment and Disease: Association or Causation? Journal of the Royal Society of Medicine 58, 295-300. https://doi.org/10.1177/003591576505800503

Hoffman, K., Lorenzo, A., Butt, C.M., Hammel, S.C., Henderson, B.B., Roman, S.A., … Sosa, J.A. (2017). Exposure to flame retardant chemicals and occurrence and severity of papillary thyroid cancer: A case-control study. Environment International 107, 235-242. https://doi.org/10.1016/j.envint.2017.06.021

International Epidemiological Association. (2014). *Dictionary of epidemiology*. *Oxford University Press*. https://doi.org/10.1093/ije/15.2.277

Last, J.M. (2001). *A Dictionary of Epidemiology*. *4th edition, Oxford, Oxford University Press*.

4.3.10.1. Question 1

Researcher A investigates the relation between parabens and breast cancer. She includes 200 women with breast cancer and 200 women without and asks all women about the use of personal care products that contain parabens in the past 10 years using a questionnaire. What is the study design that is used?

Case-control study

Cross-sectional study

Prospective cohort study

Randomized controlled trial

4.3.10.1. Question 2

Researcher B has an opinion about the advantages and disadvantages of study design that was chosen. Which of the following advantages of this design is **correct**?

There is a low chance of recall-bias.

The design is able to prove a causal relationship between the exposure and the outcome.

The design takes into account the long latency period for breast cancer.

The Relative Risk (RR) can be used in this design.

4.3.10.1. Question 3

Researcher B is involved in a prospective cohort study investigating lead exposure and ADHD in children. Prenatal exposure to lead measured in umbilical cord blood, collected at delivery. At 7 years ADHD symptoms were counted. They found that higher lead concentrations was associated with more ADHD symptoms for both boys and girls. Boys had more ADHD symptoms than girls, and higher lead levels. What kind of factor is gender?

Determinant

Confounder

Effect modifier

Outcome

4.3.10. Environmental epidemiology - II. Quantifying disease and associations

Authors: Eva Sugeng and Lily Fredrix

Reviewers: Ľubica Murínová and Raymond Niesink

Learning objectives

You should be able to

- describe measures of disease.
- calculate and interpret effect sizes fitting to the epidemiologic study design.
- describe and interpret significance level.
- describe stratification and interpret stratified data.

1. Measures of disease

Prevalence is the proportion of a population with an outcome at a certain time point (e.g. currently, 40% of the population is affected by disease Y) and can be calculated in cross-sectional studies.

Incidence concerns only new cases, and the **cumulative incidence** is the proportion of new cases in the population over a certain time span (e.g. 60% new cases of influenza per year). The (cumulative) incidence can only be calculated in prospective study designs, because the population needs to be at risk to develop the disease and therefore participants should not be affected by the disease at the start of the study.

Population Attributable Risk (PAR) is a measure to express the increase in disease in a population due to the exposure. It is calculated with this formula:

$$PAR = \frac{(incidence\ in\ the\ total\ population\ -\ incidence\ in\ the\ unexposed\ group)}{(incidence\ in\ the\ total\ population)}$$

2. Effect sizes

2.1 In case of dichotomous outcomes (disease, yes versus no)

Risk ratio or relative risk (RR) is the ratio of the **incidence** in the exposed group to the **incidence** in the unexposed group (Table 1):

$$RR = \frac{\frac{A}{A+B}}{\frac{C}{C+D}}$$

The RR can only be used in prospective designs, because it consists of probabilities of an outcome in a **population at risk**. The RR is 1 if there is no risk, <1 if there is a decreased risk, and >1 if there is an increased risk. For example, researchers find an RR of 0.8 in a hypothetical prospective cohort study on the region children live in (rural vs. urban) and the development of asthma (outcome). This means that children living in rural areas have a 0.8 lower risk to develop asthma, compared to children living in the urban areas.

Risk difference (RD) is the difference between the risks in two groups (Table 1):

$$RV = \frac{A}{A+B} - \frac{C}{C+D}$$

Odds ratio (OR) is the ratio of odds on the outcome in the exposed group to the odds of the outcome in the unexposed group (Table 1).

$$OR = \frac{\frac{A}{B}}{\frac{C}{D}}$$

The OR can be used in any study design, but is most frequently used in case-control studies. (Table 1) The OR is 1 if there is no difference in odds, >1 if there is a higher odds, and <1 if there is a lower odds. For example, researchers find an OR of 2.5 in a hypothetical case-control study on mesothelioma cancer and occupational exposure to asbestos in the past. Patients with mesothelioma cancer had 2.5 higher odds on being occupational in the past exposed to asbestos compared to the healthy controls.

The OR can also be used in terms of odds on the disease instead of the exposure, the formulae is then (Table 1):

$$OR = \frac{\frac{A}{C}}{\frac{B}{D}}$$

For example, researchers find an odds ratio of 0.9 in a cross-sectional study investigating mesothelioma cancer in builders working with asbestos comparing the use of protective clothing and masks. The builders who used protective clothing and masks had 0.9 odds on having mesothelioma cancer in comparisons to builders who did not use protective clothing and masks.

Table 1: concept table to use for calculation of the RR, RD, and OR

	Disease/outcome +	Disease/outcome -
Exposure/determinant +	A	B
Exposure/determinant -	C	D

2.2 In case of continuous outcomes (when there is a scale on which a disease can be measured, e.g. blood pressure)

Mean difference is the difference between the mean in the exposed group versus the unexposed group. This is also applicable to experimental designs with a follow-up to assess increase or decrease of the outcome after an intervention: the mean at the baseline versus the mean after the intervention. This can be standardized using the following formulae:

$$Standardized\ mean\ difference = \frac{mean\ intervention\ group\ -\ mean\ control\ group}{standard\ deviation}$$

The standard deviation (SD) is a measure of the spread of a set of values. In practice, the SD must be estimated either from the SD of the control group, or from an 'overall' value from both groups. The best-known index for effect size is Cohens 'd'. The standardized mean difference can have both a negative and a positive value (between -2.0 and +2.0). With a positive value, the beneficial effect of the intervention is shown, with a negative value, the effect is counterproductive. In general, an effect size of for example 0.8 means a large effect.

3. Statistical significance and confidence interval

Effect measurements such as the **relative risk**, the **odds ratio** and the **mean difference** are reported together with statistical **significance** and/or a **confidence interval**. Statistical significance is used to retain or reject null hypothesis. The study starts with a **null hypothesis** assumption, we assume that there is no difference between variables or groups, e.g. RR=1 or the difference in means is 0. Then the statistical test gives us the probability of getting the outcome observed (e.g. OR=2.3, or mean difference=1.5) when in fact the null hypothesis is true. If the probability is smaller than 5%, we conclude that the observation is true and we may reject the **null hypothesis**. The 5% probability corresponds to a p-value of 0.05. A p-value cut-off of $p<0.05$ is generally used, which means that p-values smaller than 0.05 are considered statistically significant.

The **95% confidence interval**. A 95% confidence interval is a range of values within which you can be 95% certain that the true mean of the population or measure of association lies. For example, in the hypothetical cross-sectional study on smoking (yes or no) and lung cancer, an OR of 2.5 was found, with an 95% CI of 1.1 to 3.5. That means, we can say with 95% certainty that the true OR lies between 1.1 and 3.5. This is regarded statistically significant, since the 1, which means no difference in odds, does not lie within the 95% CI. If researchers also studied oesophagus cancer in relation to smoking and found an OR of 1.9 with 95% CI of 0.6-2.6, this is not regarded statistically significant, since 95% CI includes 1.

4. Stratification

When two populations investigated have a different distribution of, for example, age and gender, it is often hard to compare disease frequencies among them. One way to deal with that is to analyse associations between exposure and outcome within strata (groups). This is called **stratification**. Example: a hypothetical study to investigate differences in health (**outcome**, measured with number of symptoms, such as shortness of breath while walking) between two groups of elderly, urban elderly (n=682) and rural elderly (n=143) (**determinant**). No difference between urban and rural elderly was found, however there was a difference in the number of women and men in both groups. The results for symptoms for urban and rural elderly are therefore **stratified** by gender (Table 2). Then, it appeared that male urban elderly have more symptoms than male rural elderly (p=0.01). The difference is not significant for women (p=0.07). The differences in health of elderly living an urban region are different for men and women, hence gender is an effect modifier of our association of interest.

Table 2. Number of symptoms (expressed as a percentage) for urban and rural elderly **stratified** by gender. Significant differences in bold.

number of symptoms	**Women**		**Men**	
	Urban	**Rural**	**Urban**	**Rural**
None	16.0	30.4	16.2	43.5
One	26.4	30.4	45.2	47.8
Two or more	57.6	39.1	37.8	8.7
N	125	23	74	23
p-value	0.07		**0.01**	

4.3.10.2. Question 1

Researcher B investigates the relation between parabens and breast cancer using a cross-sectional design. She includes 500 women between 55-65 years and asks them about the use of personal care products that contain parabens using a questionnaire, that divides the group into frequent users (N=267) and infrequent users (N=233). She in addition asks if the women have a breast cancer

diagnosis and finds out that 41 women have breast cancer, of these women, 30 frequently use PCP. What is a correct outcome of this study?

The Odds Ratio is 3.6.

The Population Attributable Risk is 0.4.

The incidence of breast cancer in this population is 8.2%

The prevalence of breast cancer in this population is 8.2%.

4.3.10.2. Question 2

With the statistical analysis to get the OR, Researcher A finds the following numbers: OR=2.6, 95CI: 0.9-3.5, p=0.07. What is the interpretation?

Frequent PCP users have significantly higher odds on having breast cancer.

Frequent PCP users have higher odds on breast cancer, but not significantly.

Frequent PCP users have higher risk on breast cancer, but not significantly

Frequent PCP users have significantly higher risk on breast cancer

---

4.3.10.2. Question 3

Researchers investigate an intervention to reduce phthalate exposure. The intervention group gets advice to reduce phthalates in food, personal care products and volatile household products. The control group does not get the intervention. Before and after the intervention period urine samples from the participants are collected and analyzed for phthalate metabolites. What is incorrect?

The researchers can conclude about causality, because the exposure proceeds the outcome.

The researchers can quantify the size of the difference between the intervention and control group using stratification.

The researchers can show if the intervention reduces (or increases) the exposure with a risk difference.

That there is no difference between the intervention and control group is called the null hypothesis.

4.3.11. Molecular epidemiology - I. Human biomonitoring

Author: Marja Lamoree

Reviewers: Michelle Plusquin and Adrian Covaci

Learning objectives:

You should be able to

- explain the purpose of human biomonitoring
- understand that the internal dose may come from different exposure routes
- describe the different steps in analytical methods and to clarify the specific requirements with regard to sampling, storage, sensitivity, throughput and accuracy
- clarify the role of metabolism in the distribution of samples in the human body and specify some sample matrices
- explain the role of ethics in human biomonitoring studies

Keywords: chemical analysis, human samples, exposure, ethics, cohort

Human biomonitoring

Human biomonitoring (HBM) involves the assessment of human exposure to natural and synthetic chemicals by the quantitative analysis of these compounds, their metabolites or reaction products in samples from human origin. Samples used in HBM can include blood, urine, faces, saliva, breast milk and sweat or other tissues, such as hair, nails and teeth.

The concentrations determined in human samples are a reflection of the exposure of an individual to the compounds analysed, also referred to as the internal dose. HBM data are collected to obtain insight into the population's exposure to chemicals, often with the

objective to integrate them with health data for health impact assessment in epidemiological studies. Often, specific age groups are addressed, such as neonates, toddlers, children, adolescents, adults and elderly. Human biomonitoring is an established method in occupational and environmental exposure assessment.

In several countries, HBM studies have been conducted for decades already, such as the German Environment Survey (GerES) and the National Health and Nutrition Examination Survey (NHANES) program in the United States. HBM programs may sometimes be conducted under the umbrella of the World Health Organization (WHO). Other examples are the Canadian Health Measures Survey, the Flemish Environment and Health Study and the Japan Environment and Children's Study, the latter specifically focuses on young children. Children are considered to be more at risk for the adverse health effects of early exposure to chemical pollutants, because of their rapid growth and development and their limited metabolic capacity to detoxify harmful chemicals.

Table 1. Information sources for Human Biomonitoring (HBM) programmes

Programme	Internet link
German Environment Survey (GerES)	www.umweltbundesamt.de/en/topics/health/assessing-environmentally-related-health-risks/german-environmental-survey-geres
National Health and Nutrition Examination Survey (NHANES)	https://www.cdc.gov/nchs/nhanes/index.htm
WHO	www.euro.who.int/en/data-and-evidence/environment-and-health-information-system-enhis/activities/human-biomonitoring-survey
Canadian Health Measures Survey (CHMS)	http://www23.statcan.gc.ca/imdb/p2SV...rvey&Id=148760
Japan Environment and Children's Study (JECS)	https://www.env.go.jp/en/chemi/hs/jecs/

Studies focusing on the impact of exposure to chemicals on health are conducted with the use of cohorts: groups of people that are enrolled in a certain study and volunteer to take part in the research program. Usually, apart from donating e.g. blood or urine samples, health measures, such as blood pressure, body weight, hormone levels, etc. but also data on diet, education, social background, economic status and lifestyle are collected, the latter through the use of questionnaires. A cross-sectional study aims at the acquisition of exposure and health data of the whole (volunteer) group at a defined moment, whereas in a longitudinal study follow-up studies are conducted with a certain frequency (i.e. every few years) in order to follow and evaluate the changes in exposure, describe time trends as well as study health and lifestyle on the longer term (see section on Environmental Epidemiology). To obtain sufficient statistical power to derive meaningful relationships between exposure and eventual (health) effects, the number of participants in HBM studies is often very large, i.e. ranging to 100,000 participants.

Because a lot of (sometimes sensitive) data is gathered from many individuals, ethics is an important aspect of any HBM study. Before starting a certain study involving HBM, a Medical Ethical Approval Committee needs to approve it. Applications to obtain approval require comprehensive documentation of i) the study protocol (what is exactly being investigated), ii) a statement regarding the safeguarding of the privacy and collected data of the individuals, the access of researchers to the data and the safe storage of all information and iii) an information letter for the volunteers explaining the aim of and procedures used in the study and their rights (e.g. to withdraw), so that they can give consent to be included in the study.

Chemical absorption, distribution, metabolism and excretion

Because chemicals often undergo metabolic transformation (see section on Xenobiotic metabolism and defence) after entering the body via ingestion, dermal absorption and inhalation, it is important to not only focus on the parent compound (= the compound to which the individual was exposed), but also include metabolites. Diet, socio-economic status, occupation, lifestyle and the environment all contribute to the exposure of humans, while age, gender, health status and weight of an individual define the effect of the exposure. HBM data provide an aggregation of all the different routes through which the individual was exposed. For an in-depth investigation of exposure sources, however, chemical analysis of e.g. diet (including drinking water), the indoor and outdoor environment are still necessary. Another important source of chemicals to which people are exposed in their day to day life are consumer products, such as electronics, furniture, textiles, etc., that may contain flame retardants, stain repellents, colorants and dyes, preservatives, among others.

The distribution of a chemical in the body is highly dependent on its physico-chemical properties, such as lipophilicity/hydrophilicity and persistence, while also phase I and Phase II transformation (see section on Xenobiotic metabolism and defence) play a determining role, see Figure. 1. For lipophilic compounds (e.g. section on POPs) storage occurs in fat tissue, while moderately lipophilic to hydrophilic compounds are excreted after metabolic transformation, or in unchanged form. Based on these considerations, a proper choice for sampling of the appropriate matrix can be made, i.e. some chemicals are best measured in urine, while for others blood may be better suitable.

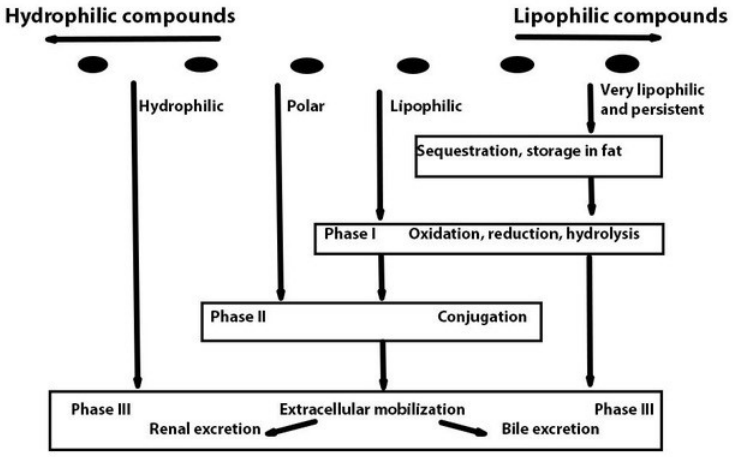


Figure 1. Distribution biotransformation of a compound (xenobiotic) in the body leading to storage or excretion.

For the design of the sampling campaign, the properties of the compounds to be analyzed should be taken into account. In case of volatility, airtight sampling containers should be used, while for light-sensitive compounds amber coloured glassware is the optimal choice.

Ideally, after collection, the samples need to be stored under the correct conditions as quickly as possible, in order the avoid degradation caused by thermal instability or by biodegradation caused by remaining enzyme activity in the sample (e.g. in blood or breast milk samples). Labeling and storage of the large quantities of samples generally included in HBM studies are important parts of the sampling campaign (see for video: https://www.youtube.com/watch?v=FQjKKvAhhjM).

Chemical analysis of human samples for exposure assessment

Typically, for the determination of the concentrations of compounds to which people are exposed and the corresponding metabolites formed in the human body, analytical techniques such as liquid and gas chromatography (LC and GC, respectively) coupled to mass spectrometry (MS) are applied. Chromatography is used for the separation of the compounds, while MS is used to detect the compounds. Prior to the analysis using LC- or GC-MS, the sample is pretreated (e.g. particles are removed) and extracted, i.e. the compounds to be analysed are concentrated in a small volume while sample matrix constituents that may interfere with the analysis (e.g. lipids, proteins) are removed, resulting in an extract that is ready to be injected onto chromatographic system.

In Figure 2 a schematic representation is given of all steps in the analytical procedure.

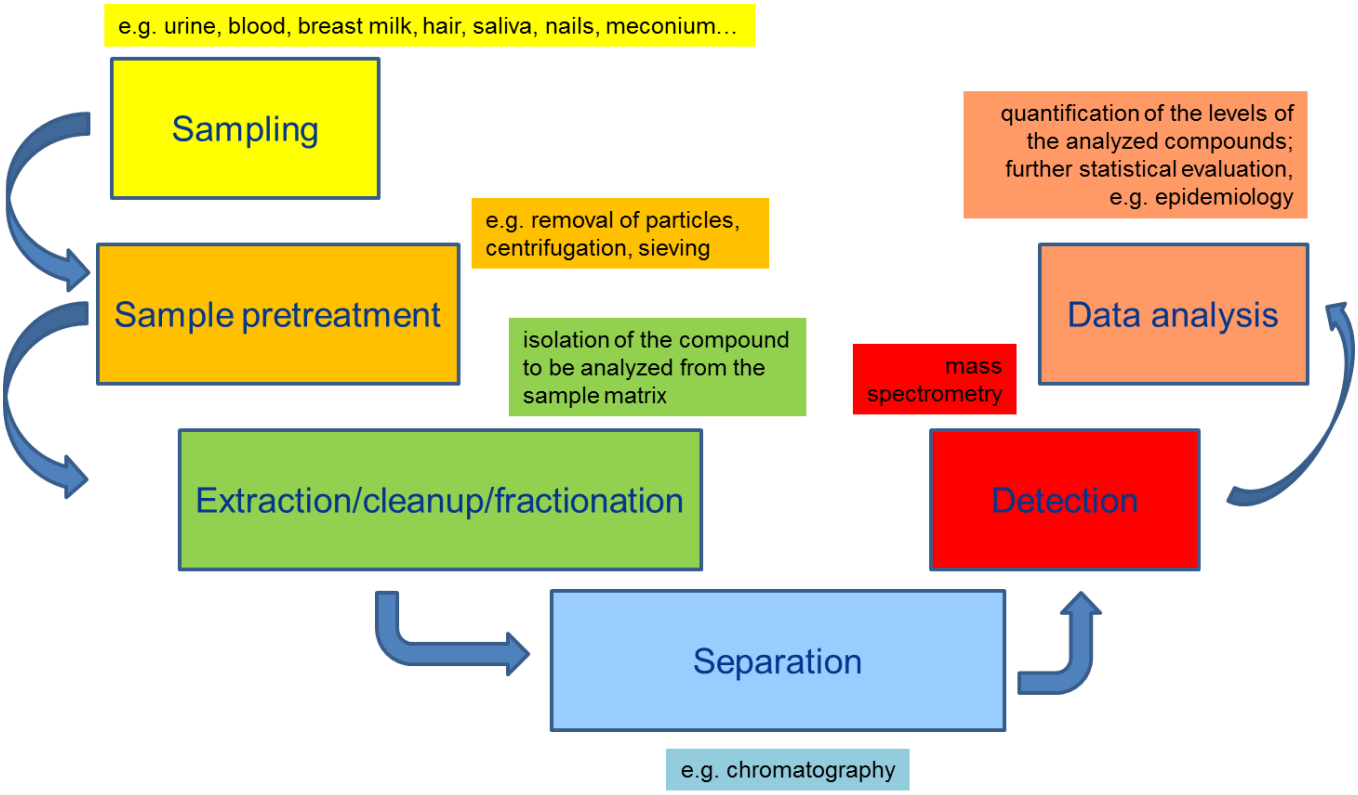


Figure 2. Schematic representation of the analytical procedure typically used for the quantitative determination of chemicals and their metabolites in human samples.

The analytical methods to quantify concentrations of chemicals in order to assess human exposure need to be of a high quality due to the specific nature of HBM studies. The compounds to be analysed are usually present in very low concentrations (i.e. in the order of pg/L for cord blood), and the sample volumes are small. For some matrices, the small sample volumes is dictated by the fact that sample availability is not unlimited, e.g. for blood. Another factor that governs the limited sample volume available is the costs that are related to the requirement of dedicated, long term storage space at conditions of -20 ^{o}C or even -80 ^{o}C to ensure sample integrity and stability.

The compounds on which HBM studies often focus are those to which we are exposed in daily life. This implies that the analytical procedure should be able to deal with contamination of the sample with the compounds to be analysed, due to the presence of the compounds in our surroundings. Higher background contamination leads to a decreased capacity to detect low concentrations, thus negatively impacting the quality of the studies. Examples of compounds that have been monitored frequently in human urine are phthalates, such as diethyl hexyl phthalate or shortly DEHP. DEHP is a chemical used in many consumer products and therefore contamination of the samples with DEHP from the surroundings severely influences the analytical measurements. One way around this is to focus on the metabolites of DEHP after Phase I or II metabolism: this guarantees that the chemical has passed the human body and has undergone a metabolic transformation, and its detection is not due to contamination from the background, which results in a more reliable exposure metric. When the analytical method is designed for the quantitative analysis of metabolites, an enzymatic step for the deconjugation of the Phase II metabolites should be included (see section on Xenobiotic metabolism and defence).

Because the generated data, i.e. the concentrations of the compounds in the human samples, are used to determine parameters like average/median exposure levels, the detection frequency of specific compounds and highest/lowest exposure levels, the accuracy of the measurements should be high. In addition, analytical methods used for HBM should be capable of high throughput, i.e. the time needed per analysis should be low, because of the large numbers of samples that are typically analysed, in the order of a hundred to a few thousand samples, depending on the study.

Summarizing, HBM data support the assessment of temporal trends and spatial patterns in human exposure, sheds light on subpopulations that are at risk and provides insight into the effectiveness of measures to reduce or even prevent adverse health

effects due to chemical exposure.

Background information:

HBM4EU project info: www.hbm4eu.eu; video: https://www.youtube.com/watch?v=DmC1v6EAeAM&feature=youtu.be.

4.3.11. Question 1

Draw a scheme of a typical analytical procedure for the quantitative determination of the concentration of a compound in a certain sample and name the different steps.

4.3.11. Question 2

What are the aims of human biomonitoring?

4.3.11. Question 3

Why is it useful to focus HBM on metabolites of chemicals and not on the original chemical, which is the case for e.g. phthalates?

4.2.3.I. Question 4

Mention 5 factors related to chemical analysis in HBM and explain their importance

---

4.3.11. Question 5

Ethics approval from a Medical Ethical Approval Committee is mandatory to carry out an HBM study. Specify what type of information should be included in a dossier in order to obtain approval.

4.3.11. Molecular epidemiology - II. The exposome and internal molecular markers

(draft)

Authors: Karen Vrijens and Michelle Plusquin

Reviewers: Frank Van Belleghem,

Learning objectives

You should be able to

- explain the concept of the exposome, including its different exposures
- understand the application of the meet-in-the-middle model in molecular epidemiological studies
- describe how different molecular markers such as gene expression, epigenetics and metabolomics can represent sensitivity to certain environmental exposure

Exposome

The exposome idea was described by Christopher Wild in 2005 as a measure of all human life-long exposures, including the process of how these exposures relate to health. An important aim of the exposome is to explain how non-genetic exposures contribute to the onset or development of important chronic disease. This concept represents the totality of exposures from three broad domains, i.e. internal, specific external and general external (Figure 1) (Wild, 2012). The internal exposome includes processes such as metabolism, endogenous circulating hormones, body morphology, physical activity, gut microbiota, inflammation, and aging. The specific external exposures include diverse agents, for example, radiation, infections, chemical contaminants and pollutants, diet, lifestyle factors (e.g., tobacco, alcohol) and medical interventions. The wider social, economic and psychological influences on the individual make up the general external exposome, including the following factors but not limited to social capital, education, financial status, psychological stress, urban-rural environment, climate, etc[1].

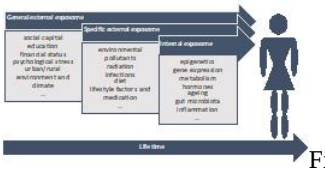


Figure 1: The exposome consists of 3 domains: the general external, the specific internal and the internal exposome.

The exposome concept clearly illustrates the complexity of the environment humans are exposed to nowadays, and how this can impact human health. There is a need for internal biomarkers of exposure (see section on Human biomonitoring) as well as biomarkers of effect, to disentangle the complex interplay between several exposures occurring potentially simultaneously and at different concentrations throughout life. Advances in biomedical sciences and molecular biology thereby collecting holistic information of epigenetics, transcriptome (see section on Gene expression), metabolome (see section on Metabolomics), etc. are at the forefront to identify biomarkers of exposure as well as of effect.

Internal molecular markers of the exposome

Meet *in the middle*model

To determine the health effect of environmental exposure, markers that can detect early changes before disease arises are essential and can be implemented in preventative medicine. These types of markers can be seen as intermediate biomarkers of effect, and their discovery relies on large-scale studies at different levels of biology (transcriptomics, genomics, metabolomics). The term "omics" refers to the quantitative measurement of global sets of molecules in biological samples using high throughput techniques (i.e. automated experiments that enable large scale repetition)[7], in combination with advanced biostatistics and bioinformatics tools[8]. Given the availability of data from high-throughput omics platforms, together with reliable measurements of external exposures, the use of omics enhances the search for markers playing a role in the biological pathway linking exposure to disease risk.

The meet-in-the-middle (MITM) concept was suggested as a way to address the challenge of identifying causal relationships linking exposures and disease outcome (Figure 2). The first step of this approach consists in the investigation of the association between exposure and biomarkers of exposure. The next step consists in the study of the relationship between (biomarkers of) exposure and intermediate omics biomarkers of early effects; and third, the relation between the disease outcome and intermediate omics biomarkers is assessed. The MITM stipulates that the causal nature of an association is reinforced if it is found in all three steps. Molecular markers that indicate susceptibility to certain environmental exposures are starting to become uncovered and can aid in targeted prevention strategies. Therefore, this approach is heavily dependent on new developments in molecular epidemiology, in which molecular biology is merged into epidemiological studies. Below, the different levels of molecular biology currently studied to identify markers of exposure and effect are discussed in detail.

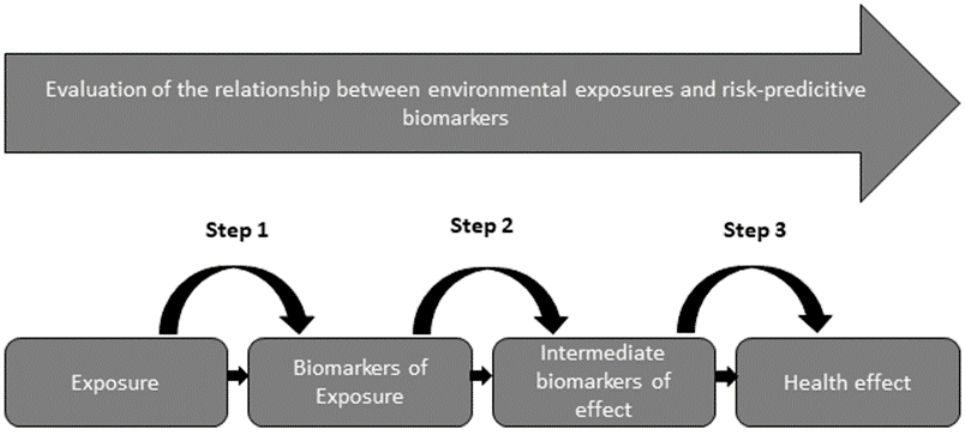


Figure adapted from Vineis & Perera (2007).

Levels

Intermediate biomarkers can be identified as measurable indicators of certain biological states at different levels of the cellular machinery, and vary in their response time, duration, site and mechanism of action. Different molecular markers might be preferred depending on the exposure(s) under study.

Gene expression

Changes at the mRNA level can be studied following a candidate approach in which mRNAs with a biological role suspected to be involved in the molecular response to a certain type of exposure (e.g. inflammatory mRNAs in the case of exposure to tobacco smoke) are selected *a priori* and measured using quantitative PCR technology or alternatively at the level of the whole genome by means of microarray analyses or Next Generation Sequencing technology. [10] Changes at the transcriptome level, are studied by analysing the totality of RNA molecules present in a cell type or sample.

Both types of studies have proven their utility in molecular epidemiology. About a decade ago the first study was published reporting on candidate gene expression profiles that were associated with exposure to diverse carcinogens[11]. Around the same time, the first studies on transcriptomics were published, including transcriptomic profiles for a dioxin-exposed population [12], in association with diesel-exhaust exposure,[13] and comparing smokers versus non-smokers both in blood [14] as well as airway epithelium cells[15]. More recently, attention has been focused on prenatal exposures in association with transcriptomic signatures, as this fits within the scope of the exposome concept. As such, transcriptomic profiles have been described in association with exposure to maternal smoking assessed in placental tissue,[16] as well as particulate matter exposure in cord blood samples[17].

Epigenetics

Epigenetics related to all heritable changes in that do not affect the DNA sequence itself directly. The most widely studied epigenetic mechanism in the field of environmental epidemiology to date is DNA methylation. DNA methylation refers to the process in which methyl-groups are added to a DNA sequence. As such, these methylation changes can alter the expression of a DNA segment without altering its sequence. DNA methylation can be studied by a candidate gene approach using a digestion-based design or, more commonly used, a bisulfite conversion followed by pyrosequencing, methylation-specific PCR or a bead array. The bisulfite treatment of DNA mediates the deamination of cytosine into uracil, and these converted residues will be read as thymine, as determined by PCR-amplification and sequencing. However, 5 mC residues are resistant to this conversion and will remain read as cytosine (Figure 3).

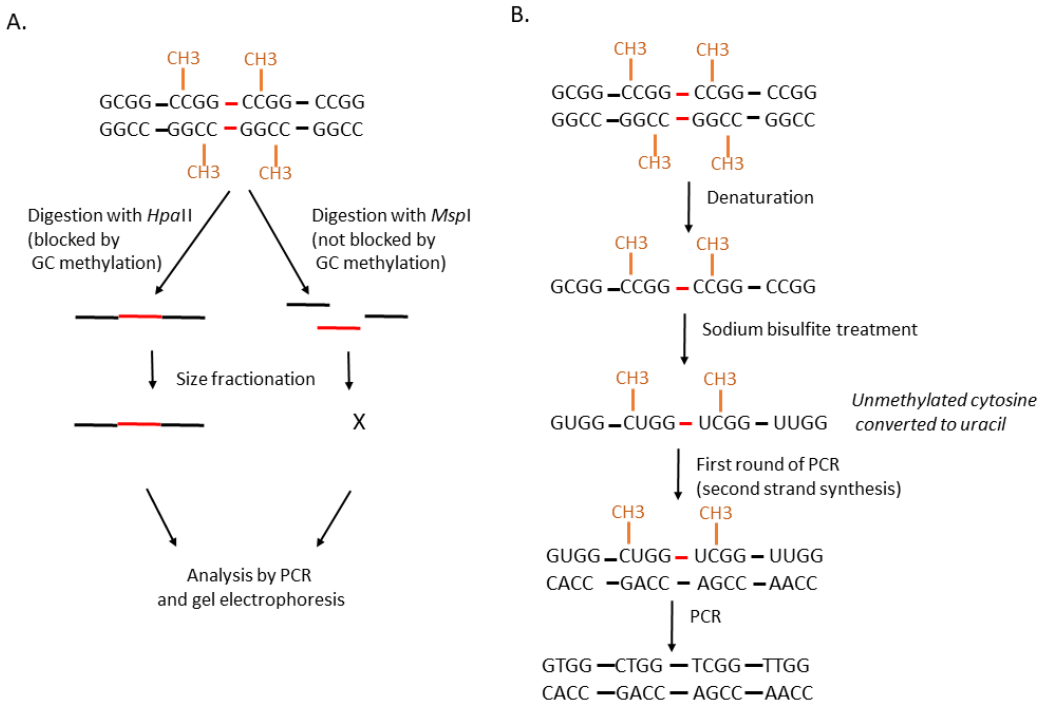


Figure 3: A. Restriction-digest based design A methylated (CH3) region of genomic DNA is digested either with two restriction enzymes, one which is blocked by GC methylation (HpaII) and one which is not(MspI). Smaller fragments are discarded (X), enriching for methylated DNA in the HpaII treated sample. B. Bisulfite-conversion of DNA. DNA is denatured and then treated with sodium bisulfite to convert unmethylated cytosine to uracil, which is converted to thymine by PCR. An important point is that following bisulfite conversion, the DNA strands are no longer complementary, and primers are designed to assay the methylation status of a specific strand.

If an untargeted approach is desirable, several strategies can be followed to obtain whole-genome methylation data, including sequencing. Epigenotyping technologies such as the human methylation BeadChips [18] generate a methylation-state-specific 'pseudo-SNP' through bisulfite conversion; therefore, translating differences in the DNA methylation patterns into sequence differences that can be analyzed using quantitative genotyping methods[19].

An interesting characteristic of DNA methylation is that it can have transgenerational effects (i.e. effects that act across multiple generations). This was first shown in a study on a population that was prenatally exposed to famine during the Dutch Hunger Winter in 1944-1945. These individuals had less DNA methylation of the imprinted gene coding for insulin-like growth factor 2 (IGF2) measured 6 decades later compared with their unexposed, same-sex siblings. The association was specific for peri-conceptional exposure (i.e. exposure during the period from before conception to early pregnancy), reinforcing that very early mammalian development is a crucial period for establishing and maintaining epigenetic marks[20].

Post-translational modifications (i.e. referring to the biochemical modification of proteins following protein biosynthesis) recently gained more attention as they are known to be induced by oxidative stress [21] (see sections on Oxidative stress) and specific inflammatory mediators [22]. Besides their function in the structure of chromatin in eukaryotic cells, histones have been shown to have toxic and pro-inflammatory activities when they are released into the extracellular space [23]. Much attention has gone to the associations between metal exposures and histone modifications,[24] although recently a first human study on the association between particulate matter exposure and histone H3 modification was published[25].

Expression of microRNAs (miRNAs are small noncoding RNAs of ∼22nt in length which are involved in the regulation of gene expression at the posttranscriptional level by degrading their target mRNAs and/or inhibiting their translation) {Ambros et al,2004} {Ambros, 2004 #324}{Ambros, 2004 #324} has also been shown to serve as a valuable marker of exposure, both candidate and

untargeted approaches have resulted in the identification of miRNA expression patterns that are associated with exposure to smoking [26], particulate matter [27], and chemicals such as polychlorinated biphenyls (PCBs) [28].

Metabolomics

Metabolomics have been proposed as a valuable approach to address the challenges of the exposome. Metabolomics, the study of metabolism at the whole-body level, involves assessment of the entire repertoire of small molecule metabolic products present in a biological sample. Unlike genes, transcripts and proteins, metabolites are not encoded in the genome. They are also chemically diverse, consisting of carbohydrates, amino acids, lipids, nucleotides and more. Humans are expected to contain a few thousand metabolites, including those they make themselves as well as nutrients and pollutants from their environment and substances produced by microbes in the gut. The study of metabolomics increases knowledge on the interactions between gene and protein expression, and the environment[29]. Metabolomics can be a biomarker of effect of environmental exposure as it allows for the full characterization of biochemical changes that occur during xenobiotic metabolism (see Section on Xenobiotic metabolism and defence). Recent technological developments have allowed downscaling the sample volume necessary for the analysis of the full metabolome, allowing for the assessment of system-wide metabolic changes that occur as a result of an exposure or in conjunction with a health outcome [30]. As for all discussed biomarkers, both targeted metabolomics, in which specific metabolites are measured in order to characterize a pathway of interest, as well as untargeted metabolomic approaches are available. Among "omics" methodologies, metabolomics interrogates the levels of a relatively lower number of features as there are about 2900 known human metabolites versus ~30,000 genes. Therefore it has strong statistical power compared to transcriptome-wide and genome-wide studies [31]. Metabolomics is, therefore, a potentially sensitive method for identifying biochemical effects of external stressors. Even though the developing field of "environmental metabolomics" seeks to employ metabolomic methodologies to characterize the effects of environmental exposures on organism function and health, the relationship between most of the chemicals and their effects on the human metabolome have not yet been studied.

Challenges

Limitations of molecular epidemiological studies include the difficulty to obtain samples to study, the need for large study populations to identify significant relations between exposure and the biomarker, the need for complex statistical methods to analyse the data. To circumvent the issue of sample collection, much effort has been focused on eliminating the need for blood or serum samples by utilizing saliva samples, buccal cells or nail clippings to read out molecular markers. Although these samples can be easily collected in a non-invasive manner, care must be taken to prove that these samples indeed accurately reflect the body's response to exposure rather than a local effect. For DNA methylation, it has been shown this is heavily dependent on the locus under study. For certain CpG sites the correlation in methylation levels is much higher than for other sites [32]. For those sites that do not correlate well across tissues, it has furthermore been demonstrated that DNA methylation levels can differ in their associations with clinical outcomes [33], so care must be taken in epidemiological study design to overcome these issues.

4.3.11. Question 6

Describe the 3 components of the exposome

4.3.11. Question 7

Explain how using the meet in the middle model biomarkers can be identified

---

4.3.11. Question 8

Give 3 examples of molecular (bio)markers of environmental exposure.

4.3.12. Gene expression

Author: Nico M. van Straalen

Reviewers: Dick Roelofs, Dave Spurgeon

Learning objectives:

You should be able to

- provide an overview of the various "omics" approaches (genomics, transcriptomics, proteomics and metabolomics) deployed in environmental toxicology.
- describe the practicalities of transcriptomics, how a transcription profile is generated and analysed.
- indicate the advantages and disadvantages of the use of genome-wide gene expression in environmental toxicology.
- develop an idea on how transcriptomics might be integrated into risk assessment of chemicals.

Keywords: genomics, transcriptomics, proteomics, metabolomics, risk assessment

Synopsis

Low-dose exposure to toxicants induces biochemical changes in an organism, which aim to maintain homoeostasis of the internal environment and to prevent damage. One aspect of these changes is a high abundance of transcripts of biotransformation enzymes, oxidative stress defence enzymes, heat shock proteins and many proteins related to the cellular stress response. Such defence mechanisms are often highly inducible, that is, their activity is greatly upregulated in response to a toxicant. It is also known that most of the stress responses are specific to the type of toxicant. This principle may be reversed: if an upregulated stress response is observed, this implies that the organism is exposed to a certain stress factor; the nature of the stress factor may even be derived from the transcription profile. For this reason, microarrays, RNA sequencing or other techniques of transcriptome analysis, have been applied in a large variety of contexts, both in laboratory experiments and in field surveys. These studies suggest that transcriptomics scores high on (in decreasing order) (1) rapidity, (2) specificity, and (3) sensitivity. While the promises of genomics applications in environmental toxicology are high, most of the applications are in mode-of-action studies rather than in risk assessment.

Introduction

No organism is defenceless against environmental toxicants. Even at exposures below phenotypically visible no-effect levels a host of physiological and biochemical defence mechanisms are already active and contribute to the organism's homeostasis. These regulatory mechanisms often involve upregulation of defence mechanisms such as oxidative stress defence, biotransformation (xenobiotic metabolism), heat shock responses, induction of metal-binding proteins, hypoxia response, repair of DNA damage, etc. At the same time downregulation is observed for energy metabolism and functions related to growth and reproduction. In addition to these targeted regulatory mechanisms targeting, there are usually a lot of secondary effects and dysfunctional changes arising from damage. A comprehensive overview of all these adjustments can be obtained from analysis of the transcriptome.

In this module we will review the various approaches adopted in "omics", with an emphasis on transcriptomics. "Omics" is a container term comprising five different activities. Table 1 provides a list of these approaches and their possible contribution to environmental toxicology. Genomics and transcriptomics deal with DNA and mRNA sequencing, proteomics relies on mass spectrometry while metabolomics involves a variety of separation and detection techniques, depending on the class of compounds analysed. The various approaches gain strength when applied jointly. For example proteomics analysis is much more insightful if it can be linked to an annotated genome sequence and metabolism studies can profit greatly from transcription profiles that include the enzymes responsible for metabolic reactions. Systems biology aims to integrate the different approaches using mathematical models. However, it is fair to say that the correlation between responses at the different levels is often rather poor. Upregulation of a transcript does not always imply more protein, more protein can be generated without transcriptional upregulation and the concentration of a metabolite is not always correlated with upregulation of the enzymes supposed to produce it. In this module we will focus on transcriptomics only. Metabolomics is dealt with in a separate section.

Table 1. *Overview of the various "omics" approaches*

Term	Description	Relevance for environmental toxicology
Genomics	Genome sequencing and assembly, comparison of genomes, phylogenetics, evolutionary analysis	Explanation of species and lineage differences in susceptibility from the structure of targets and metabolic potential, relationship between toxicology, evolution and ecology
Transcriptomics	Genome-wide transcriptome (mRNA) analysis, gene expression profiling	Target and metabolism expression indicating activity, analysis of modes of action, diagnosis of substance-specific effects, early warning instrument for risk assessment
Proteomics	Analysis of the protein complement of the cell or tissue	Systemic metabolism and detoxification, diagnosis of physiological status, long-term or permanent effects
Metabolomics	Analysis of all metabolites from a certain class, pathway analysis	Functional read-out of the physiological state of a cell or tissue
Systems biology	Integration of the various "omics" approaches, network analysis, modelling	Understanding of coherent responses, extrapolation to whole-body phenotypic responses

Transcriptomics analysis

The aim of transcriptomics in environmental toxicology is to gain a complete overview of all changes in mRNA abundance in a cell or tissue as a function of exposure to environmental chemicals. This is usually done in the following sequence of steps:

1. Exposure of organisms to an environmental toxicant, including a range of concentrations, time-points, etc., depending on the objectives of the experiment.
2. Isolation of total RNA from individuals or a sample of pooled individuals. The number of biological replicates is determined at this stage, by the number of independent RNA isolations, not by technical replication further on in the procedure.
3. Reverse transcription. mRNAs are transcribed to cDNA using the enzyme reverse transcriptase that initiates at the polyA tail of mRNAs. Because ribosomal RNA lacks a poly(A)tail they are (in principle) not transcribed to cDNA. This is followed by size selection and sometimes labelling of cDNAs with barcodes to facilitate sequencing.
4. Sequencing of the cDNA pool and transcriptome assembly. The assembly preferably makes use of a reference genome for the species.If no reference genome is available, the transcriptome is assembled *de novo*, which requires a greater sequencing depth and usually ends in many incomplete transcripts. A variety of corrections are applied to equalize effects of total RNA yield, library size, sequencing depth, gene length, etc.
5. Gene expression analysis and estimation of fold regulation. This is done, in principle, by counting the normalized number of transcripts per gene for every gene in the genome, for each of the different conditions to which the organism was exposed. The response per gene is expressed as fold regulation, by expressing the transcripts relative to a standard or control condition. Tests are conducted to separate significant changes from noise.
6. Annotation and assessment of pathways and functions as influenced by exposure. An integrative picture is developed, taking all evidence together, of the functional changes in the organism.

In the recent past, step 4 was done by microarray hybridization rather than by direct sequencing. In this technique two pools of cDNA (e.g. a control and a treatment) are hybridized to a large number of probes fixed onto a small glass plate. The probes are designed to represent the complete gene complement of the organism. Positive hybridization signals are taken as evidence for upregulated gene expression. Microarray hybridization arose in the years 1995-2005 but has now been largely overtaken by ultrafast and high-throughput next generation sequencing methods, however, due to cost-efficiency, relative simplicity of bioinformatics analysis, and standardization of the assessed genes it is still often used.

We illustrate the principles of transcriptomics analysis and the kind of data analysis that follows it, by an example from the work by Bundy *et al*. (2008). These authors exposed earthworms (*Lumbricus rubellus*) to soils experimentally amended with copper, quite a toxic element for earthworms. The copper-induced transcriptome was surveyed using a custom-made microarray and metabolic profiles were established using NMR (nuclear magnetic resonance) spectroscopy. From the 8,209 probes on the microarray, 329 showed a significant alteration of expression under the influence of copper. The data were plotted in a "heat map" diagram (Figures 1A and 1B), providing a quick overview of upregulated and downregulated genes. The expression profiles were also analysed in reduced dimensionality using principal component analysis (PCA). This showed that the profiles varied considerably with treatment. Especially the highest and the penultimate highest exposures generated a profile very different from the control (see Figure 1C). The genes could be allocated to four clusters, (1) genes upregulated by copper over all exposures (Figure 1D), (2) genes downregulated by copper (see Figure 1E), (3) genes upregulated by low exposures but unaffected at higher exposures (see Figure 1F), and (4) genes upregulated by low exposure but downregulated by higher concentrations (see Figure 1G). Analysis of gene identity combined with metabolite analysis suggested that the changes were due to an effect of copper on mitochondrial respiration, reducing the amount of energy generated by oxidative phosphorylation. This mechanism was underlying the reduction of body-growth observed on the phenotypic level.

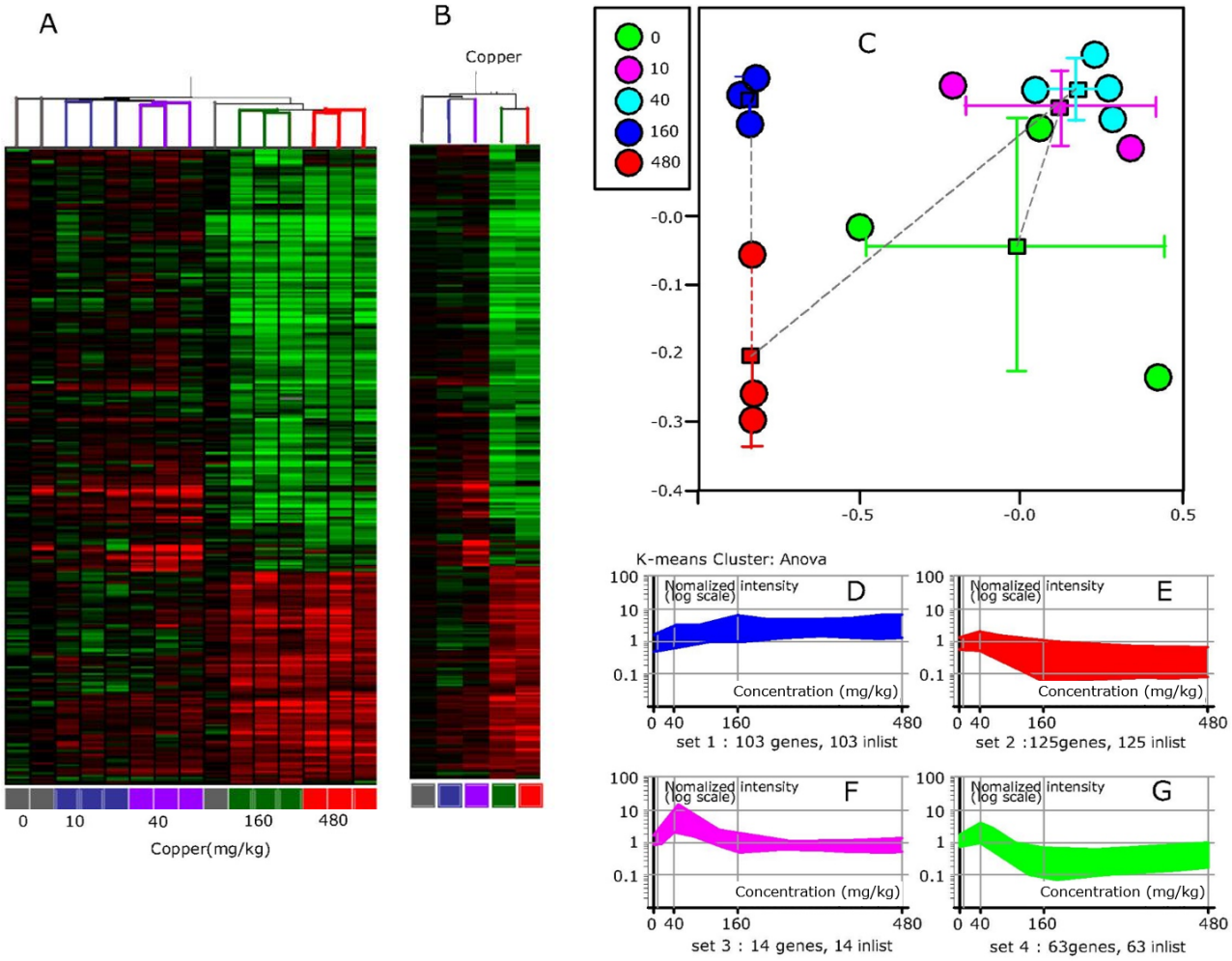


Figure 1. Example of a transcriptomics analysis aiming to understand copper toxicity to earthworms. A "heat map" of individual replicates (four in each of five copper treatments). Expression is indicated for each of the 329 differentially expressed genes (arranged from top to bottom) in red (downregulated) or green (upregulated). A cluster analysis showing the similarities is indicated above the profiles. B. The same data, but with the four replicates per copper treatment joined. The data show that at 40 mg/kg of copper in soil some of the earthworm's genes are starting to be downregulated, while at 160 mg/kg and 480 mg/kg significant upregulation and downregulation is occurring. C Principal Component Analysis of the changes in expression profile. The multivariate expression profile is reduced to two dimensions and the position of each replicate is indicated by a single point in the biplot; the confidence interval over four replicates of each copper treatment is indicated by horizontal and vertical bars. The

profiles of the different copper treatments (joined by a dashed line) differ significantly from each other. D, E, F, and G. Classification of the 329 genes in four groups according to their responses to copper (plotted on the horizontal axis). Redrawn from Bundy et al. (2008) by Wilma Ijzerman.

Omics in risk assessment

How could omics-technology, especially transcriptomics, contribute to risk assessment of chemicals? Three possible advantages have been put forward:

1. Gene expression analysis is *rapid*. Gene regulation takes place on a time-scale of hours and results can be obtained within a few days. This compares very favourably with traditional toxicity testing (*Daphnia*, 48 hours, *Folsomia*, 28 days).
2. Gene expression is *specific*. Because a transcription profile involves hundreds to thousands of endpoints (genes), the information content is potentially very large. By comparing a new profile generated by an unknown compound, to a trained data set, the compound can usually be identified quite precisely.
3. Gene expression is *sensitive*. Because gene regulation is among the very first biochemical responses in an organism, it is expected to respond to lower dosages, at which whole-body parameters such as survival, growth and reproduction are not yet responding.

Among these advantages, the second one (specificity) has shown to be the most consistent and possibly brings the largest advantage. This can be illustrated by a study by Dom *et al.* (2012) in which gene expression profiles were generated for *Daphnia magna* exposed to different alcohols and chlorinated anilines (Figure 2).

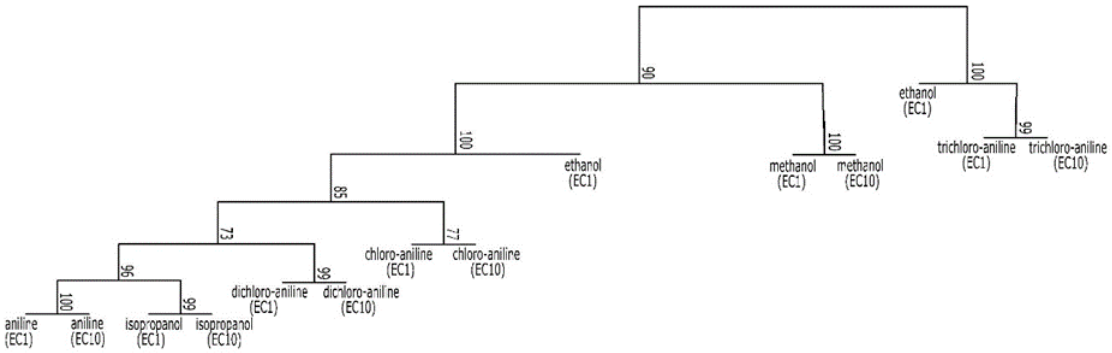


Figure 2.
Clustered gene expression profiles of Daphnia magna exposed to seven different compounds. Replicates exposed to the same compound are clustered together, except for ethanol. The first split separates exposures that at the EC10 level (reproduction) did not show any effects on growth and energy reserves (right) and exposures that caused significant such effects (left). Reproduced from Dom et al. (2012) by Wilma IJzerman.

The profiles of replicates exposed to the same compound were always clustered together, except in one case (ethanol), showing that gene expression is quite specific to the compound. It is possible to reverse this argument: from the gene expression profile the compound causing it can be deduced. In addition, the example cited showed that the first separation in the cluster analysis was between exposures that did and did not affect energy reserves and growth. So the gene expression profiles are not only indicative of the compound, but also of the type of effects expected.

The claim of rapidity also proved true, however, the advantage of rapidity is not always borne out. It may be an issue when quick decisions are crucial (evaluating a truck loaded with suspect contaminated soil, deciding whether to discharge a certain waste stream into a lake yes or no), but for regular risk assessment procedures it proved to be less of an advantage than sometimes expected. Finally, greater sensitivity of gene expression, in the sense of lower no-observed effect concentrations than classical endpoints is a potential advantage, but proves to be less spectacular in practice. However, there are clear examples in which exposures below phenotypic effect levels were shown to induce gene expression responses, indicating that the organism was able to compensate any negative effects by adjusting its biochemistry.

Another strategy regarding the use of gene expression in risk assessment is not to focus on genome-wide transcriptomes but on selected biomarker genes. In this strategy, gene expressions are aimed for that show (1) consistent dose-dependency, (2) responses over a wide range of contaminants, and (3) correlations with biological damage. For example, De Boer *et al.* (2015) analysed a

composite data set including experiments with six heavy metals, six chlorinated anilines, tetrachlorobenzene, phenanthrene, diclofenac and isothiocyanate, all previously used in standardized experiments with the soil-living collembolan, *Folsomia candida*. Across all treatments a selection of 61 genes was made, that were responsive in all cases and fulfilled the three criteria listed above. Some of these marker genes showed a very good and reproducible dose-related response to soil contamination. Two biomarkers are shown in Figure 3. This experiment, designed to diagnose a field soil with complex unknown contamination, clearly demonstrated the presence of *Cyp*-inducing organic toxicants.

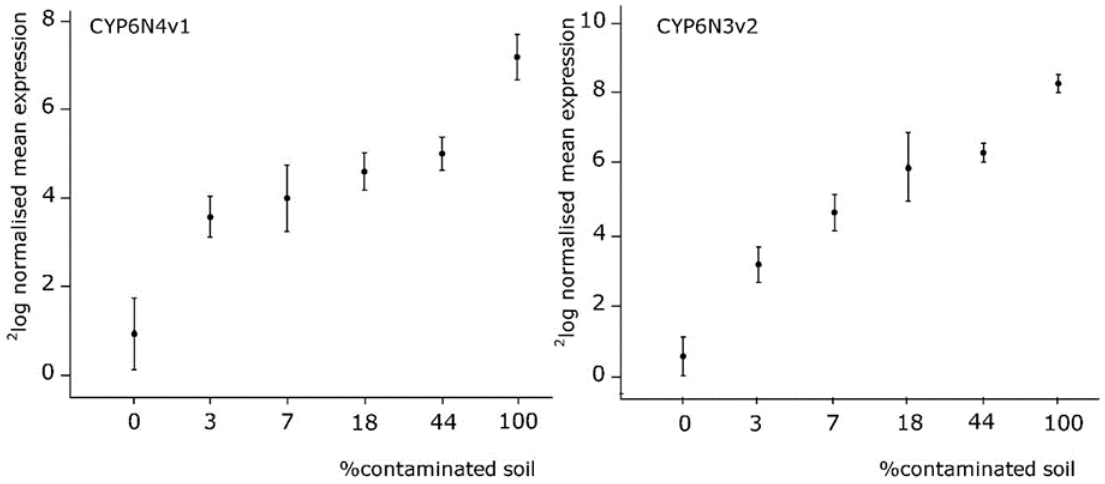


Figure 3. Showing gene expression, relative to control expression, for two selected biomarker genes (encoding cytochroms P450 phase I biotransformation enzymes) in the genome of the soil-living collembolan Folsomia candida, in response to the concentration of contaminated field soil spiked-in into a clean soil at different rates. Reproduced from Roelofs et al. (2012) by Wilma IJzerman.

Of course there are also disadvantages associated with transcriptomics in environmental toxicology, for example:

1. Gene expression analysis requires a knowledge-intensive infrastructure, including a high level of expertise for some of the bioinformatics analyses. Also, adequate molecular laboratory facilities are needed; some techniques are quite expensive.
2. Gene expression analysis is most fruitful when species are used that are backed up by adequate genomic resources, especially a well annotated genome assembly, although this is becoming less of a problem with increasing availability of genomic resources.
3. The relationship between gene expression and ecologically relevant variables such as growth and reproduction of the animal is not always clear.

Conclusions

Gene expression analysis has come to occupy a designated niche in environmental toxicology since about 2005. It is a field highly driven by technology, and shows continuous change over the last years. It may significantly contribute to risk assessment in the context of mode of action studies and as a source of designated biomarker techniques. Finally, transcriptomics data are very suitable to feed into information regarding key events, important biochemical alterations that are causally linked up to the level of the phenotype to form an adverse outcome pathway. We refer to the section on Adverse outcome pathways for further reading.

References

Bundy, J.G., Sidhu, J.K., Rana, F., Spurgeon, D.J., Svendsen, C., Wren, J.F., Stürzenbaum, S.R., Morgan, A.J., Kille, P. (2008). "Systems toxicology" approach identifies coordinated metabolic responses to copper in a terrestrial non-model invertebrate, the earthworm *Lumbricus rubellus*. BMC Biology 6, 25.

De Boer, T.E., Janssens, T.K.S., Legler, J., Van Straalen, N.M., Roelofs, D. (2015). Combined transcriptomics analysis for classification of adverse effects as a potential end point in effect based screening. Environmental Science and Technology 49, 14274-14281.

Dom, N., Vergauwen, L., Vandenbrouck, T., Jansen, M., Blust, R., Knapen, D. (2012). Physiological and molecular effect assessment versus physico-chemistry based mode of action schemes: *Daphnia magna* exposed to narcotics and polar narcotics. Environmental Science and Technology 46, 10-18.

Gibson, G., Muse, S.V. (2002). *A Primer of Genome Science*. Sinauer Associates Inc., Sunderland.

Gibson, G. (2008). The environmental contribution to gene expression profiles. *Nature* Reviews Genetics 9, 575-581.

Roelofs, D., De Boer, M., Agamennone, V., Bouchier, P., Legler, J., Van Straalen, N. (2012). Functional environmental genomics of a municipal landfill soil. Frontiers in Genetics 3, 85.

Van Straalen, N.M., Feder, M.E. (2012). Ecological and evolutionary functional genomics - how can it contribute to the risk assessment of chemicals? Environmental Science & Technology 46, 3-9.

Van Straalen, N.M., Roelofs, D. (2008). Genomics technology for assessing soil pollution. Journal of Biology 7, 19.

4.3.12. Question 1

Define: (1) genomics, (2) transcriptomics (3) proteomics, and (4) metabolomics, and indicate which of these approaches is most useful in the risk assessment of chemicals.

4.3.12. Question 2

Specify three advantages and at least one disadvantage of the use of genome-wide gene expression as a tool in environmental risk assessment.

4.3.13. Metabolomics

Author: Pim E.G. Leonards

Reviewers: Nico van Straalen, Drew Ekman

Learning objectives:

You should be able to:

- understands the basics of metabolomics and how metabolomics can be used.
- describe the basic principles of metabolomics analysis, and how a metabolic profile is generated and analysed.
- describe the differences between targeted and untargeted metabolomics and how each is used in environmental toxicology.
- develop an idea on how metabolomics might be integrated into hazard and risk assessments of chemicals.

Keywords: Metabolomics, metabolome, environmental metabolomics, application areas of metabolomics, targeted and untargeted metabolomics, metabolomics analysis and workflow

Introduction

Metabolomics is the systematic study of small organic molecules (<1000 Da) that are intermediates and products formed in cells and biofluids due to metabolic processes. A great variety of small molecules result from the interaction between genes, proteins and metabolites. The primary types of small organic molecules studied are endogenous metabolites (i.e., those that occur naturally in the cell) such as sugars, amino acids, neurotransmitters, hormones, vitamins, and fatty acids. The total number of endogenous metabolites in an organism is still under study but is estimated to be in the thousands. However, this number varies considerably between species and cell types. For instance, brain cells contain relative high levels of neurotransmitters and lipids, nevertheless the levels between different types of brain tissues can vary largely. Metabolites are working in a network, e.g. citric acid cycle, by the conversion of molecules by enzymes. The turnover time of the metabolites is regulated by the enzymes present and the amount of metabolites present.

The field of metabolomics is relatively new compared to genomics, with the first draft of the human metabolome available in 2007. However, the field has grown rapidly since that time due to its recognized ability to reflect molecular changes most closely associated with an organism's phenotype. Indeed, in comparison to other 'omics approaches (e.g., transcriptomics), metabolites are the downstream results from the action of genes and proteins and, as such, provide a direct link with the phenotype (Figure 1). The metabolic status of an organism is directly related to its function (e.g. energetic, oxidative, endocrine, and reproductive status) and phenotype, and is, therefore, uniquely suitable to relate chemical stress to the health status of organisms. Moreover, transcriptomics and proteomics, the identification of metabolites does not require the existence of gene sequences, making it particularly useful for the those species which lack a sequenced genome.

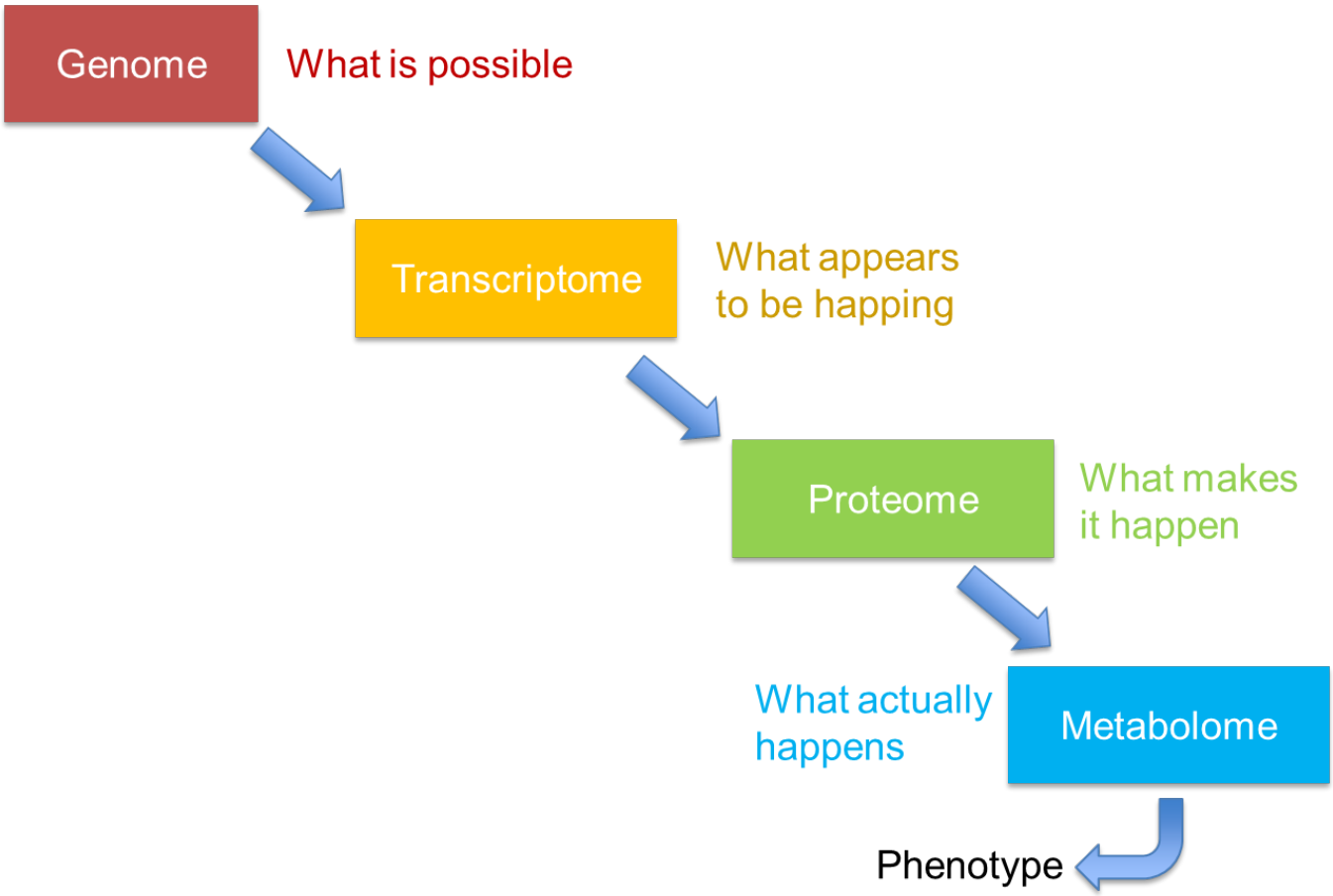


Figure 1: Cascade of different omics fields.

Definitions

The *complete set* of small molecules in a biological system (e.g. cells, body fluids, tissues, organism) is called the *metabolome* (Table 1). The term metabolomics was introduced by Oliver et al. (1998) who described it "as the complete set of metabolites/low molecular weight compounds which is context dependent, varying according to the physiology, development or pathological state of the cell, tissue, organ or organism". This quote highlights the observation that the levels of metabolites can vary due to internal as well as external factors, including stress resulting from exposure to environmental contaminants. This has resulted in the emergence and growth of the field of *environmental metabolomics* which is based on the application of metabolomics, to biological systems that are exposed to environmental contaminants and other relevant stressors (e.g., temperature). In addition to endogenous metabolites, some metabolomic studies also measure changes in the biotransformation of environmental contaminants, food additives, or drugs in cells, the collection of which has been termed the *xenometabolome*.

Table 1: Definitions of metabolomics.

Term	Definition	**Relevance for environmental toxicology**
Metabolomics	Analysis of small organic molecules (<1000 Da) in biological systems (e.g. cell, tissue, organism)	Functional read-out of the physiological state of a cell or tissue and directly related to the phenotype
Metabolome	Measurement of the *complete set* of small molecules in a biological system	Discovery of affected metabolic pathways due to contaminant exposure
Environmental metabolomics	Metabolomics analysis in biological systems that are exposed to environmental stress, such as the exposure to environmental contaminants	Metabolomics focused on environmental contaminant exposure to study for instance the mechanism of toxicity or to find a biomarker of exposure or effect
Xenometabolome	Metabolites formed from the biotransformation of environmental contaminants, food additives, or drugs	Understanding the metabolism of the target contaminant
Targeted metabolomics	Analysis of a pre-selected set of metabolites in a biological system	Focus on the effects of environmental contaminants on specific metabolic pathways
Untargeted metabolomics	Analysis of all detectable (i.e., not preselected) of metabolites in a biological system	Discovery-based analysis of the metabolic pathways affected by environmental contaminant exposure

Environmental Metabolomics Analysis

The development and successful application of metabolomics relies heavily on i) currently available analytical techniques that measure metabolites in cells, tissues, and organisms, ii) the identification of the chemical structures of the metabolites, and iii) characterisation of the metabolic variability within cells, tissues, and organisms.

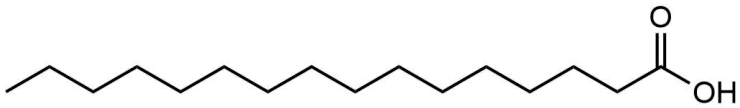
The aim of metabolomics analysis in environmental toxicology can be:

- to focus on changes in the abundances of **specific** metabolites in a biological system

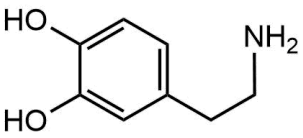
after environmental contaminant exposure: *targeted* metabolomics

- to provide a "complete" overview of changes in abundances of **all detectable** metabolites in a biological system after environmental contaminant exposure: *untargeted* metabolomics

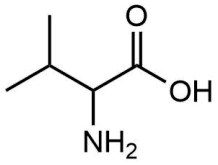
In *targeted* metabolomics a limited number of pre-selected metabolites (typically 1-100) are quantitatively analysed (e.g. nmol dopamine/g tissue). For example, metabolites in the neurotransmitter biosynthetic pathway could be targeted to assess exposures to pesticides. Targeting specific metabolites in this way typically allows for their detection at low concentrations with high accuracy. Conversely, in untargeted metabolomics the aim is to detect as many metabolites as possible, regardless of their identities so as to assess as much of the metabolome as possible. The largest challenge for untargeted metabolomics is the identification (annotation) of the chemical structures of the detected metabolites. There is currently no single analytical method able to detect all metabolites in a sample, and therefore a combination of different analytical techniques are used to detect the metabolome. Different techniques are required due to the wide range of physical-chemical properties of the metabolites. The variety of chemical structures of metabolites are shown in Figure 2. Metabolites can be grouped in classes such as fatty acids (the classes are given in brackets in Figure 2), and within a class different metabolites can be found.



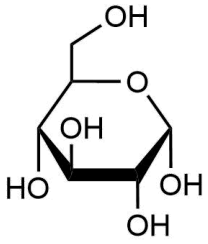
Palmitic acid (fatty acid)



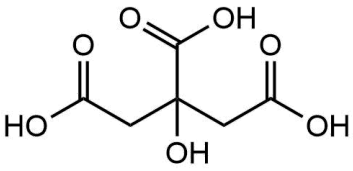
Dopamine (neurotransmitter)



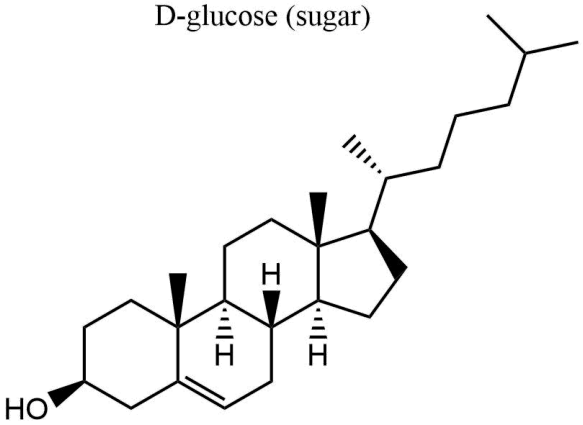
valine (amino acid)



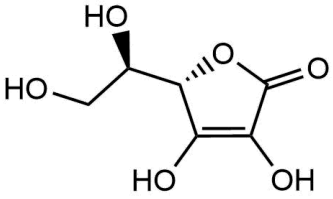
D-glucose (sugar)



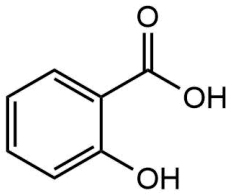
Citric acid (organic acid)



Cholesterol (sterol lipid)



Vitamin C



Salicylic acid (phytohormone)

Figure 2: Examples of the chemical structures of several commonly detected metabolites. Metabolite classes are indicated in brackets. Drawn by Steven Droge.

A general workflow of environmental metabolomics analysis uses the following steps:

1. of the organism or cells to an environmental contaminant. An unexposed control group must also be included. The exposures often include the use of various concentrations, time-points, etc., depending on the objectives of the study.
2. **Sample collection** of the relevant biological material (e.g. cell, tissue, organism). It is important that the collection be done as quickly as possible so as to quench any further metabolism. Typically ice cold solvents are used.
3. of the metabolites from the cell, tissue or organisms by a two-step extraction using a combination of polar (e.g. water/methanol) and apolar (e.g. chloroform) extraction solvents.
4. of the polar and apolar fractions using liquid chromatography (LC) or gas chromatography (GC) combined with mass spectrometry (MS), or by nuclear magnetic resonance (NMR) spectroscopy. The analytical tool(s) used will depend on the

metabolites under consideration and whether a targeted or untargeted approach is required.

5. **Metabolite detection** (targeted or untargeted analysis).
    - Targeted metabolomics - a specific set of pre-selected metabolites are detectedand their concentrations are determined using authentic standards.
    - Untargeted metabolomics - a list of all detectable metabolites measured by MS or NMR response, and their intensities is collected. Various techniques are then used to determine the identities of those metabolites that change due to the exposure (see step 7 below).
6. **Statistical analysis** using univariate and multivariate statistics to calculate the difference between the exposure and the control groups. The fold change (fold increase or decrease of the metabolite levels) between an exposure and control group are determined.
7. For untargeted metabolomics only the chemical structure of the statistically significant metabolites are identified. The **identification** of the chemical structure of the metabolite can be based on the molecular weight, isotope patterns, elemental compositions, mass spectrometry fragmentation patterns, etc. Mass spectrometry libraries are used for the identification to match the above parameters in the samples with the data in libraries.
8. **Data processing**: Identification of the metabolic pathways influenced by the chemical exposure. Integrative picture of the molecular and functional level of an organism due to chemical exposure. Understand the relationship between the chemical exposure, molecular pathway changes and the observed toxicity. Or to identify potential biomarkers of exposure or effect.

Box: Analytical tools for metabolomics analysis

The most frequently used analytical tools for measuring metabolites are mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy. MS is an analytical tool that generates ions of molecules and then measures their mass-to-charge ratios. This information can be used to generate a "molecular finger print" for each molecule, and based on this finger print metabolites can be identified. Chromatography is typically used to separate the different metabolites of a mixture found in a sample before it enters the mass spectrometer. Two main chromatography techniques are used in metabolomics: liquid chromatography and gas chromatography. Due to its high sensitivity, MS is able to measure a large number of different metabolites simultaneously. Moreover, when coupled with a separation method such as chromatography, MS can detect and identify thousands of metabolites.

Mass spectrometry is much more sensitive than NMR, and it can detect a large range of different types of metabolites with different physical-chemical properties. NMR is less sensitive and can therefore detect a lower number of metabolites (typically 50-200). The advantages of NMR are the minimum amount of sample handling, the reproducibility of the measurements (due to high precision), and it is easier to quantify the levels of metabolites. In addition, NMR is a non-destructive technique such that a sample can often be used for further analyses after the data has been acquired.

Application of environmental metabolomics

Metabolomics has been widely used in drug discovery and medical sciences. More recently, metabolomics is being incorporated into environmental studies, an emerging field of research called environmental metabolomics. Environmental metabolomics is used mainly in five application domains (Table 2). Arguably the most commonly used application is for studying the mechanism of toxicity/mode of action (MoA) of contaminants. However, many studies have identified select metabolites that show promise for use as biomarkers of exposure or effect. As a result of its strength in identifying response fingerprints, metabolomics is also finding use in the regulatory toxicology field particularly for read-across studies. This application is particularly useful for rapidly screening contaminants for toxicity. Metabolomics can be also be used in dose-response studies (benchmark dosing) to derive a point of departure (POD). This is especially interesting in regulatory chemical risk assessment.

Currently, the field of systems toxicology is explored by combining data from different omics field (e.g. transcriptomics, proteomics, metabolomics) to improve our understanding of the relationship between the different omics, chemical exposure, and toxicity, and to better understand the mechanism of toxicity/ MoA.

Table 2: Application areas of metabolomics in environmental toxicology.

Application area	Description
Mechanism of toxicity/ Mode of action (MoA)	Using metabolomics to understand at the molecular level the pathways that are affected by exposure to environmental contaminants. Discovery of the mode of action of chemicals. In an adverse outcome pathway (AOP)Discovery metabolomics is used to identify the key events (KE), by linking chemical exposure at the molecular level to functional endpoints (e.g. reproduction, behaviour).
Biomarker discovery	Identification of metabolites that can be used as convenient (i.e., easy and inexpensive to measure) indicators of exposure or effect.
Read-across	In regulatory toxicology, metabolomics is used in read-across studies to provide information on the similarity of the responses between chemicals. This approach is useful to identify more environmentally toxic chemicals.
Point of departure	Metabolomics can be used in dose-response studies (benchmark dosing) to derive a point of departure (POD). This is especially interesting in regulatory chemical risk assessment. This application is currently not used yet.
Systems toxicology	Combination approach of different omics (e.g. transcriptomics, proteomics, metabolomics) to improve our understanding of the relationship between the different omics and chemical exposure, and to better understand the mechanism of toxicity/ MoA.

As an illustration of the mechanism of toxicity/mode of action application, Bundy et al. (2008) used NMR-based metabolomics to study earthworms (*Lumbricus rubellus*) exposed to various concentrations of copper in soil (0, 10, 40, 160, 480 mg copper / kg soil). They performed both transcriptomic and metabolomics studies. Both polar (sugars, amino acid, etc.) and apolar (lipids) metabolites were analysed, and fold changes relative to the control group were determined. For example, differences in the fold changes of lipid metabolites (e.g. fatty acids, triacylglycerol) as a function of copper concentration are shown as a "heatmap" in Figure 3A. Clearly the highest dose group (480 mg/kg) has a very different lipid metabolite pattern than the other groups. The polar metabolite data was analysed using principal component analysis (PCA) , a multivariate statistical tool that reduces the number of dimensions of the data. The PCA score plot shown in Figure 3B reveals that the largest differences in metabolite profiles exist between: the control and low dose (10 mg Cu/kg) groups, the 40 mg Cu/kg and 160 mg Cu/kg groups, and the highest dose (480 mg Cu/kg) group (Figure 3B). These separations indicate that the metabolite patterns in these groups were different as a result of the different copper exposures. Some of the metabolites were up- and some were down-regulated due to the copper exposure (two examples given in Figures 3c and 3D). The metabolite data were also combined with gene expression data in a systems toxicology application. This combined analysis showed that the copper exposures led to disruption of energy metabolism, particularly with regard to effects on the mitochondria and oxidative phosphorylation. Bundy et al. associated this effect on energy metabolism with a reduced growth rate of the earthworms. This study effectively showed that metabolomics can be used to understand the metabolite pathways that are affected by copper exposure and are closely linked to phenotypic changes (i.e., reduced growth rate). The transcriptome data collected simultaneously were in good accordance with the metabolome patterns, supporting Bundy et al.'s hypothesis that simultaneous measurement of the transcriptomes and metabolome can be used to validate the findings of both approaches, and in turn the value of "systems toxicology".

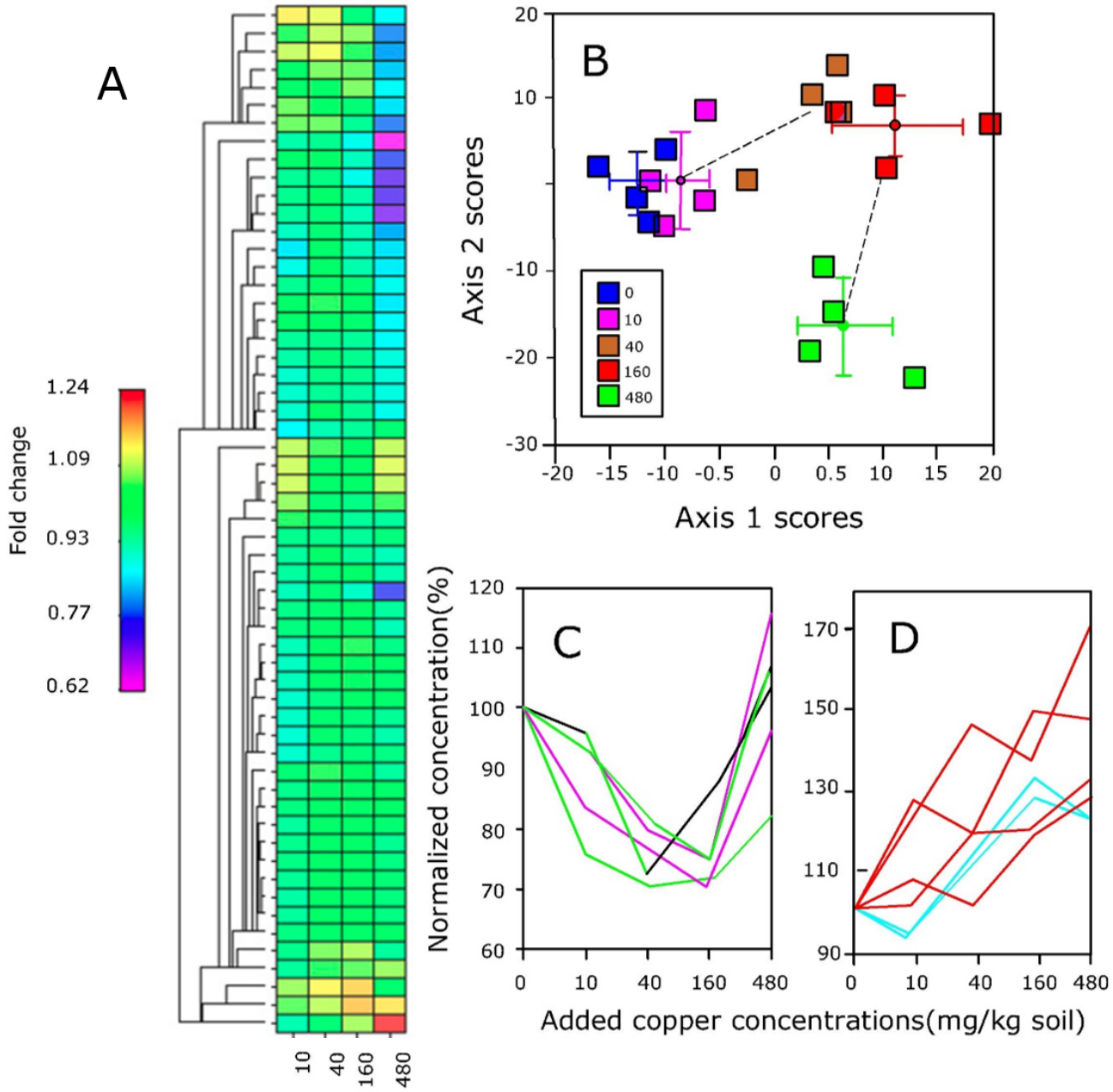


Figure 3: Example of metabolite analysis with NMR to understand the mechanism of toxicity of copper to earthworms (Bundy et al., 2008). A: Heatmap, showing the fold changes of lipid metabolites at different exposure concentrations of copper (10, 40, 160, 480 mg/kg copper in soil) and controls. B: Principal component analysis (PCA) of the polar metabolite patterns of the exposure groups. The higher dose group (480 mg/kg soil) is separate from the medium dose (40 mg Cu/kg and 160 mg Cu/kg) groups and the control and the lowest dose groups (0 mg Cu/kg and 10 mg Cu/kg soil) indicating that the metabolite patterns in these groups are different and are affected by the copper exposure. C: Down and upregulation of lipophilic amino acids (blue: aliphatics, red: aromatics). D: Upregulation of cell-membrane-related metabolites (black = betaine, glycine, HEFS, phosphoethanolamine, red: myo-inositol, scyllo-inositol). Redrawn from Bundy et al. (2008) by Wilma IJzerman.

Challenges in metabolomics

Several challenges currently exist in the field of metabolomics. From a biological perspective, metabolism is a dynamic process and therefore very time-sensitive. Taking samples at different time-points during development of an organism, or throughout a chemical exposure can result in quite different metabolite patterns. Sample handling and storage can also be challenging as some metabolites are very unstable during sample collection and sample treatment. From an analytical perspective, metabolites possess a

wide range of physico-chemical properties and occur in highly varying concentrations such that capturing the widest portion of the metabolome requires analysis with more than one analytical technique. However, the largest challenge is arguably the identification of the chemical structure of unknown metabolites. Even with state-of-the-art analytical techniques only a fraction of the unknown metabolites can be confidently identified.

Conclusions

Metabolomics is a relative new field in toxicology, but is rapidly increasing our understanding of the biochemical pathways affected by exposure to environmental contaminants, and in turn their mechanisms of action. Linking the changed molecular pathways due to the contaminant exposure to phenotypical changes of the organisms is an area of great interest. Continual advances in state-of-the-art analytical tools for metabolite detection and identification will continue to this trend and expand the utility of environmental metabolomics for prioritizing contaminants. However, a number of challenges remain for the widespread use of metabolomics in regulatory toxicology. Fortunately, recent growth in international interest to address these challenges is underway, and is making great strides in a variety of applications.

References

Bundy, J.G., Sidhu, J.K., Rana, F., Spurgeon, D.J., Svendsen, C., Wren, J.F., Sturzenbaum, S.R., Morgan, A.J., Kille, P. (2008). 'Systems toxicology' approach identifies coordinated metabolic responses to copper in a terrestrial non-model invertebrate, the earthworm *Lumbricus rubellus*. BMC Biology, 6(25), 1-21.

Bundy, J.G., Matthew P. Davey, M.P., Viant, M.R. (2009). Environmental metabolomics: a critical review and future perspectives. Metabolomics, 5, 3-21.

Johnson, C.H., Ivanisevic, J., Siuzdak, G. (2016). Metabolomics: beyond biomarkers and towards mechanisms. Nature Reviews, Molecular and Cellular Biology 17, 451-459.

4.3.13. Question 1

What type of molecules are measured with metabolomics?

Proteins

Small molecules

Genes

Polymers

4.3.13. Question 2

What are typical application areas of environmental metabolomics?

4.3.13. Question 3

Give four main elements of the workflow of metabolomics, and describe two in more detail?

This page titled 2.3.4.3: Toxicity Testing is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.3: Toxicity Testing** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.4.4: Increasing ecological realism in toxicity testing

4.4. Increasing ecological realism in toxicity testing

Author: Michiel Kraak

Reviewer: Kees van Gestel

Learning objectives:

You should be able to

- argue the need to increase ecological realism in single-species toxicity tests.
- list the consecutive steps to increase ecological realism in single-species toxicity tests.

Keywords: single-species toxicity tests, mixture toxicity, multistress, chronic toxicity, multigeneration effects, ecological realism.

Introduction

The vast majority of single-species toxicity tests reported in the literature concerns acute or short-term exposures to individual chemicals, in which mortality is often the only endpoint. This is in sharp contrast with the actual situation at contaminated sites, where organisms may be exposed to relatively low levels of mixtures of contaminants under suboptimal conditions for their entire life span. Hence there is an urgent need to increase ecological realism in single-species toxicity tests by addressing sublethal endpoints, mixture toxicity, multistress effects, chronic toxicity and multigeneration effects.

Increasing ecological realism in single-species toxicity tests

Mortality is a crude parameter representing the response of organi

sms to relatively high and therefore often environmentally irrelevant toxicant concentrations. At much lower and environmentally more relevant toxicant concentrations, organisms may suffer from a wide variety of sublethal effects. Hence, the first step to gain ecological realism in single-species toxicity tests is to address sublethal endpoints instead of, or in addition to mortality (Figure 1). Yet, given the short exposure time in acute toxicity tests it is difficult to assess other endpoints than mortality. Photosynthesis of plants and behaviour of animals are elegant, sensitive and rapidly responding endpoints that can be incorporated into short-term toxicity tests to enhance their ecological realism (see section on Endpoints).

Since organisms are often exposed to relatively low levels of contaminants for their entire life span, the next step to increase ecological realism in single-species toxicity tests is to increase exposure time by performing chronic experiments (Figure 1) (see section on Chronic toxicity). Moreover, in chronic toxicity tests a wide variety of sublethal endpoints can be assessed in addition to mortality, the most common ones being growth and reproduction (see to section on Endpoints). Given the relatively short duration of the life cycle of many invertebrates and unicellular organisms like bacteria and algae, it would be relevant to prolong the exposure time even further, by exposing the test organisms for their entire life span, so from the egg or juvenile phase till adulthood including their reproductive performance, or for several generations, assessing multigeneration effects (Figure 1) (see section on Multigeneration effects).

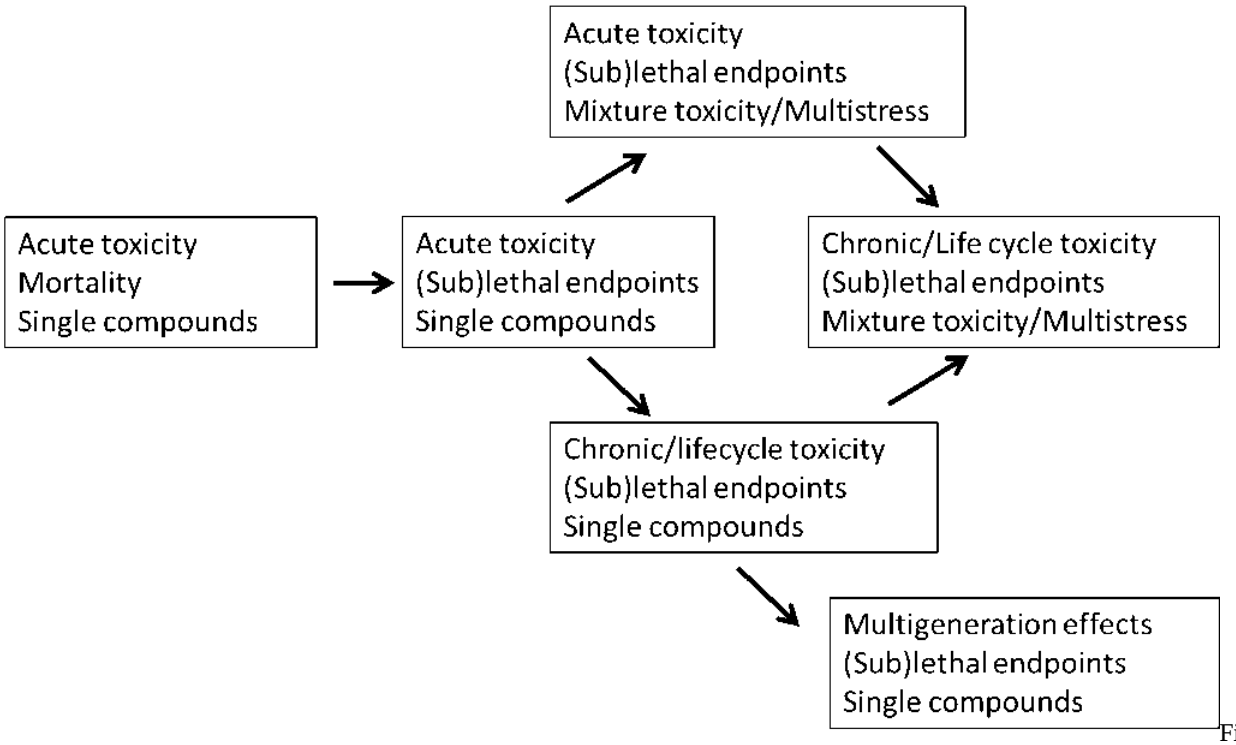


Figure 1.

Consecutive steps of increasing ecological realism in single-species toxicity tests.

In contaminated environments, organisms are generally exposed to a wide variety of toxicants under variable and sub-optimal conditions. To further gain ecological realism, mixture toxicity and multistress scenarios should thus be considered (figure 1) (see sections on Mixture toxicity and Multistress). The highest ecological relevance of laboratory toxicity tests may be achieved by addressing the above mentioned issues all together in one type of experiment, chronic mixture toxicity tests assessing sublethal endpoints. Yet, even nowadays such studies remain scarce.

Another way of increasing ecological realism of toxicity testing is by moving towards multispecies test systems that allow for assessing the impacts of chemicals and other stressors on species interactions within communities (see chapter 5 on Population, community and ecosystem ecotoxicology).

4.4. Question 1

Argue the need to increase ecological realism in single-species toxicity tests.

4.4 Question 2

List the consecutive steps to increase ecological realism in single-species toxicity tests.

4.4.1. Mixture toxicity

Authors: Michiel Kraak & Kees van Gestel

Reviewer: Thomas Backhaus

Learning objectives:

You should be able to

· explain the concepts involved in mixture toxicity testing, including Concentration Addition and Response Addition.

· design mixture toxicity experiments and to understand how the toxicity of (equitoxic) toxicant mixtures is assessed.

· interpret the results of mixture toxicity experiments and to understand the meaning of Concentration Addition, Response Addition, as well as antagonism and synergism as deviations from Concentration Addition and Response Addition.

Key words: Mixture toxicity, TU summation, equitoxicity, Concentration Addition, Response Addition, Independent Action

Introduction

In contaminated environments, organisms are generally exposed to complex mixtures of toxicants. Hence, there is an urgent need for assessing their joint toxic effects. In theory, there are four classes of joint effects of compounds in a mixture as depicted in Figure 1.

Four classes of joint effects	***No interaction (additive)***	***Interaction (non-additive)***
Similar action	Simple similar action/ Concentration Addition	Complex similar action
Dissimilar action	Independent action/ Response Addition	Dependent action

Figure 1. *The four classes of joint effects of compounds in a mixture, as proposed by Hewlett and Plackett (1959).*

Simple similar action & Concentration Addition

The most simple case concerns compounds that share the same mode of action and do not interact (Figure 1 upper left panel: simple similar action). This holds for compounds acting on the same biological pathway, affecting strictly the same molecular target. Hence, the only difference is the relative potency of the compounds. In this case Concentration Addition is taken as the starting point, following the Toxic Unit (TU) approach. This approach expresses the toxic potency of a chemical as TU, which is calculated for each compound in the mixture as:

$$Toxic \quad Unit = \frac{c}{EC_x}$$

with c = the concentration of the compound in the mixture, and EC_x = the concentration of the compound where the measured endpoint is affected by X % compared to the non-exposed control. Next, the toxic potency of the mixture is calculated as the sum of the TUs of the individual compounds:

$$TU(mixture) = sumTU(compounds) \quad = \sum \frac{c_i}{EC_{x,1}}$$

Imagine that the EC_{50} of compound A is 300 µg.L^{-1} and that of compound B 60 µg.L^{-1}. In a mixture of A+B 30 µg.L^{-1} A and 30 µg.L^{-1} B are added. These concentrations represent 30/300 = 0.1 TU of A and 30/60 = 0.5 TU of B. Hence, the mixture consists of 0.1 + 0.5 = 0.6 TU. Yet, the two compounds in this mixture are not represented at equal toxic strength, since this specific mixture is dominated by compound B. To compose mixtures in which the compounds are represented at equal toxic strength, the equitoxicity concept is applied:

1 Equitoxic TU A+B = 0.5 TU A + 0.5 TU B

1 Equitoxic TU A+B = 150 µg.L^{-1} A + 30 µg.L^{-1} B

As in traditional concentration-response relationships, survival or a sublethal endpoint is plotted against the mixture concentration from which the EC_{50} value and the corresponding 95% confidence limits can be derived (see section on Concentration-response relationships). If the upper and lower 95% confidence limits of the EC_{50} value of the mixture include 1 TU, the EC_{50} of the mixture does not differ from 1 TU and the toxicity of the compounds in the mixture is indeed concentration additive (Figure 2).

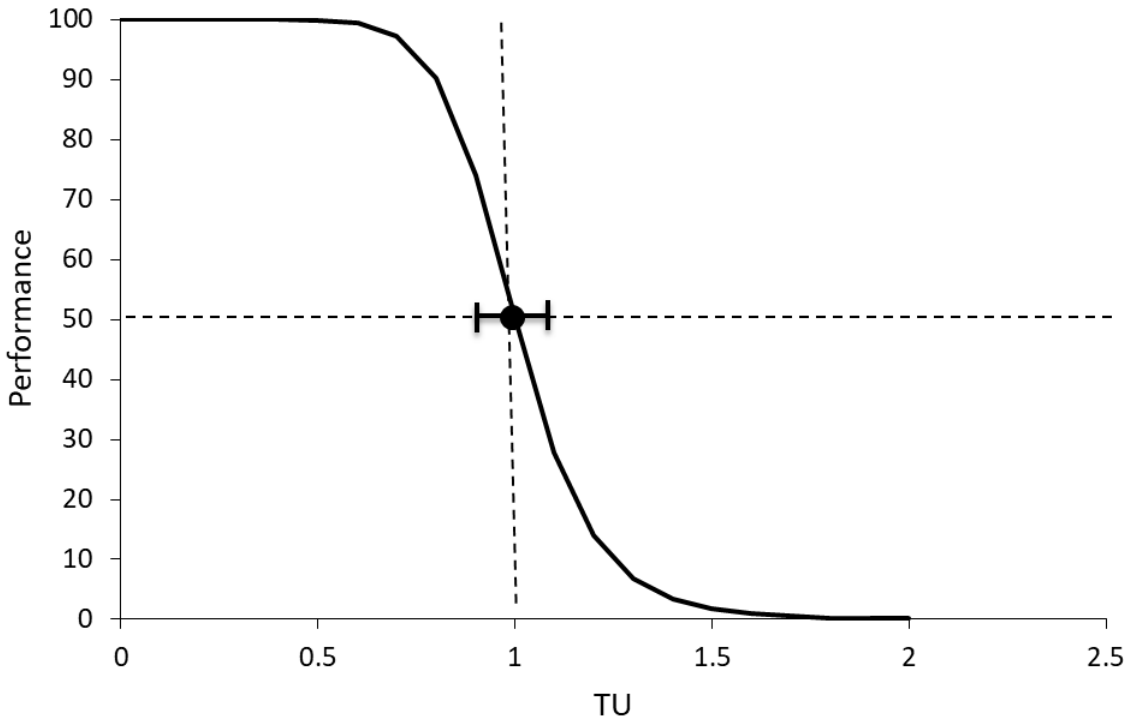


Figure 2. Concentration-response relationship for a mixture in which the toxicants have a concentration additive effect. The Y-axis shows the performance of the test organisms, e.g. their survival, reproduction or other endpoint measured. The horizontal dotted line represents the 50% effect level, the vertical dotted line represents 1 Toxic Unit (TU). The black dot represents the experimental EC_{50} value of the mixture with the 95% confidence limits.

An experiment appealing to the imagination was performed by Deneer et al. (1988), who tested a mixture of 50 narcotic compounds (see section on Toxicodynamics and Molecular interactions) and observed perfect concentration addition, even when the individual compounds were present at only 0.25% (0.0025 TU) of their EC_{50}. This showed in particular that narcotic compounds present at concentrations way below their no effect level still contribute to the joint toxicity of a mixture (Deneer et al., 1988). This was also shown for metals (Kraak et al., 1999). This is alarming, since even nowadays environmental legislation is still based on a compound-by-compound approach. The study by Deneer et al. (1988) also clearly demonstrated the logistical challenges of mixture toxicity testing. Since for composing equitoxic mixtures the EC_{50} values of the individual compounds need to be known, testing an equitoxic mixture of 50 compounds requires 51 toxicity tests: 50 individual compounds and 1 mixture.

Independent Action & Response Addition

When chemicals have a different mode of action, act on different targets, but still contribute to the same biological endpoint, the mixture is expected to behave according to Response Addition (also termed Independent Action; Figure 1, lower left panel). Such a situation would occur, for example, if one compound inhibits photosynthesis, and a second one inhibits DNA-replication, but both inhibit the growth of an exposed algal population. To calculate the effect of a mixture of compounds with different modes of action, Response Addition is applied as follows: The probability that a compound, at the concentration at which it is present in the mixture, exerts a toxic effect (scaled from 0 to 1), differs per compound and the cumulative effect of the mixture is the result of combining these probabilities, according to:

$$E_{(mix)} = E_{(A)} + E_{(B)} - E_{(A)}E_{(B)}$$

Where $E_{(mix)}$ is the fraction affected by the mixture, and $E_{(A)}$ and $E_{(B)}$ are the fractions affected by the individual compounds A and B at the concentrations at which they occur in the mixture. In fact, this equation sums the fraction affected by compound A and the fraction affected by compound B at the concentration at which they are present in the mixture, and then corrects for the fact that the

fraction already affected by chemical A cannot be affected again by chemical B (or vice versa). The latter part of the equation is needed to account for the fact that the chemicals act independent from each other. This is visualised in Figure 3.

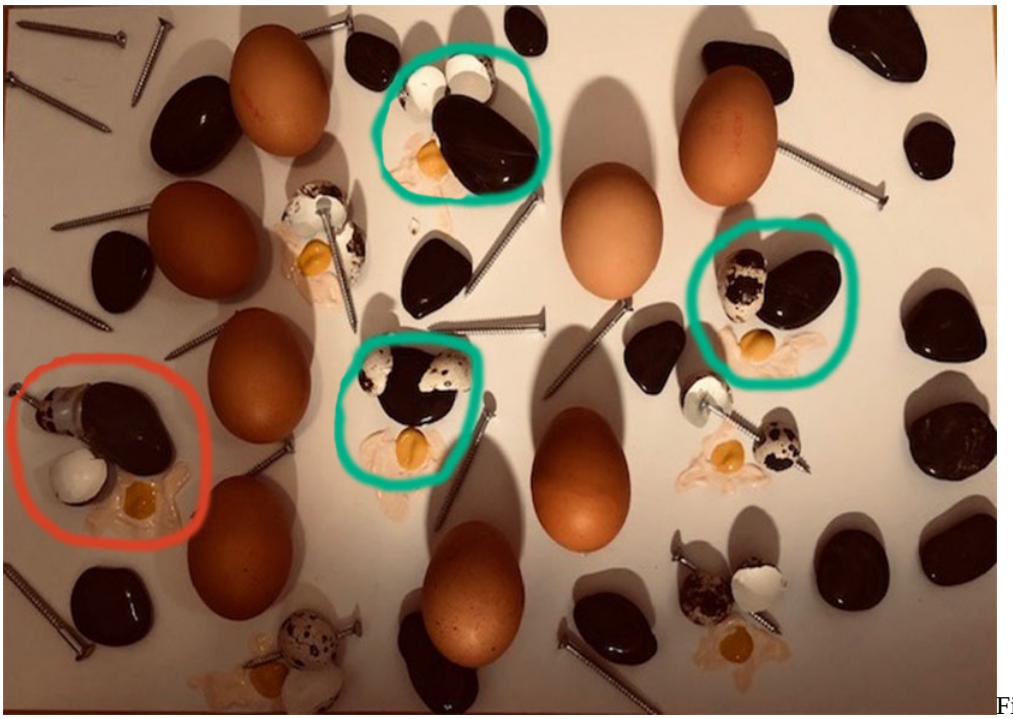


Figure 3. Illustration of stressors acting independent of each other, using the example given by Berenbaum (1981). Subsequently a handful of nails and a handful of pebbles are thrown to a collection of eggs. The nails break 5 eggs, and these 5 eggs broken by the nails cannot be broken again. The pebbles could break 4 eggs, but 1 egg was already broken by the nails. Hence, the pebbles break 3 additional eggs.

The equation: $E_{(mix)} = E_{(A)} + E_{(B)} - E_{(A)}E_{(B)}$

can be rewritten as: $1-E_{(mix)} = (1-E_A)*(1-E_B)$

This means that the probability of not being affected by the mixture ($1-E_{(mix)}$) is the product of the probabilities of not being affected by (the specific concentrations of) compound A and compound B. At the EC_{50}, both the affected and the unaffected fraction are 50%, hence $(1-E_A)*(1-E_B) = 0.5$. If both compounds equally contribute to the effect of the mixture, $(1-E_A) = (1-E_B)$ and thus $(1-E_{A \text{ or } B})^2 = 0.5$, so both $(1-E_A)$ and $(1-E_B)$ equal $\sqrt{0.5} = 0.71$. Since the probability of not being affected is 0.71 for compound A and compound B, the probability of being affected is 0.29. Thus at the EC_{50} of a mixture of two compounds acting according to Independent Action, both compounds should be present at a concentration equalling their EC_{29}.

Interactions between the compounds in a mixture

Concentration Addition as well as Response Addition both assume that the compounds in a mixture do not interact (see Figure 1). However, in reality, such interactions can occur in all four steps of the toxic action of a mixture. The *first step* concerns chemical and physicochemical interactions. Compounds in the environment may interact, affecting each other's bioavailability. For instance, excess of Zn causes Cd to be more available in the soil solution as a result of competition for the same binding sites. The *second step* involves physiological interactions during uptake by an organism, influencing the toxicokinetics of the compounds, for example by competition for uptake sites at the cell membrane. The *third step* refers to the internal processing of the compounds, e.g. involving effects on each other's biotransformation or detoxification (toxicokinetics). The *fourth step* concerns interactions at the target site(s), i.e. the toxicodynamics during the actual intoxication process. The typical whole organism responses that are recorded in many ecotoxicity tests integrate the last three types of interactions, resulting in deviations from the toxicity predictions from Concentration Addition and Response Addition.

Deviations from Concentration Addition

If the EC_{50} of the mixture is higher than 1 TU and the lower 95% confidence limit is also above 1 TU, the toxicity of the compounds in the mixture is less than concentration additive, as more of the mixture is needed than anticipated to cause 50% effect

(Figure 4, blue line; antagonism). Correspondingly, if the EC_{50} of the mixture is lower than 1 TU and the upper 95% confidence limit is also below 1 TU, the toxicity of the compounds in the mixture is more than concentration additive (Figure 4, red line; synergism).

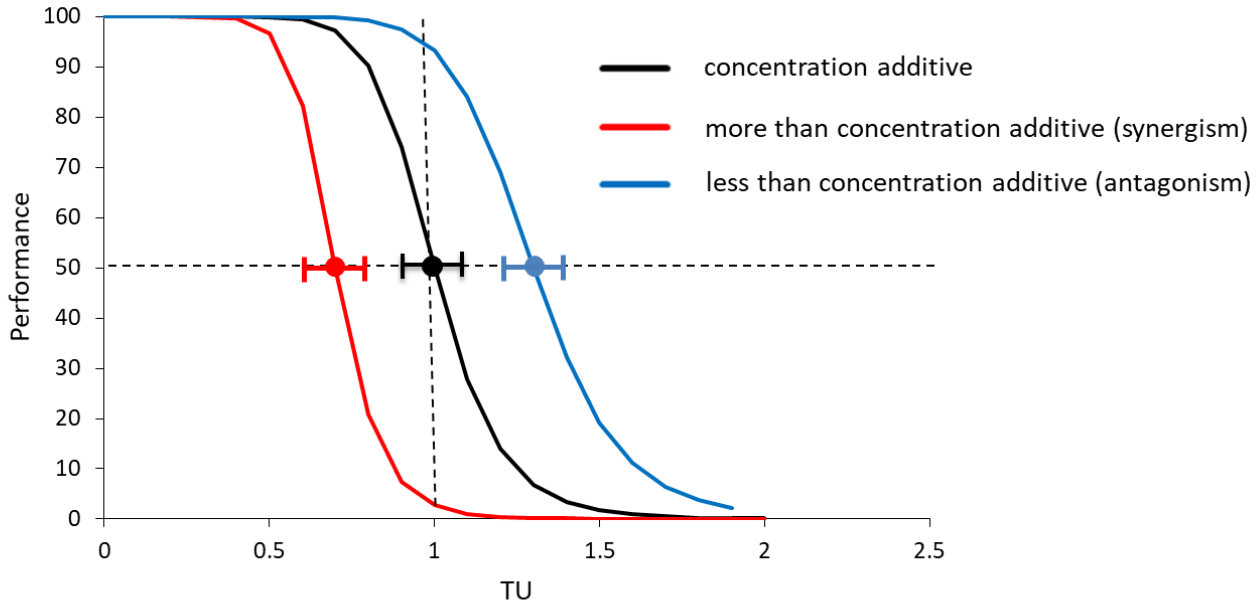


Figure 4. Concentration-response relationships for mixtures in which the toxicants have a less than concentration additive effect (blue line), a concentration additive effect (black line) and a more than concentration additive effect (red line). The Y-axis shows the performance of the test organisms, e.g. their survival, reproduction or other endpoint measured. The horizontal dotted line represents the 50% effect level, the vertical dotted line represents 1 Toxic Unit (TU). The coloured dots represent the EC_{50} values with the corresponding 95% confidence limits.

When the toxicity of a mixture is more than concentration additive, the compounds enhance each other's toxicity. When the toxicity of a mixture is less than concentration additive, the compounds reduce each other's toxicity. Both types of deviation from additivity can have two different reasons: 1. The compounds have the same mode of action, but do interact (Figure 1, upper right panel: complex similar action). 2. The compounds have different modes of actions (Independent action/Response Addition; Figure 1, lower left panel).

Concentration-response surfaces and isoboles

Elaborating on Figure 4, concentration-response relationships for mixtures can also be presented as multi-dimensional figures, with different axes for the concentration of each of the chemicals included in the mixture (Figure 5A). In case of a mixture of two chemicals, such a dose-response surface can be shown in a two-dimensional plane using isoboles. Figure 5B shows isoboles for a mixture of two chemicals, under different assumptions for interactions according to Concentration Addition. If the interaction between the two compounds decreases the toxicity of the mixture, this is referred to as *antagonism* (Figure 5B, blue line). If the interaction between the two compounds increases the toxicity of the mixture, this is referred to as *synergism* (Figure 5B, red line). Thus both antagonism and synergism are terms to describe deviations from Concentration Addition due to interaction between the compounds. Yet, antagonism in relation to Concentration Addition (less than concentration additive; Figure 5B blue line) can simply be caused by the compounds behaving according to Response Addition, and not behaving antagonistically.

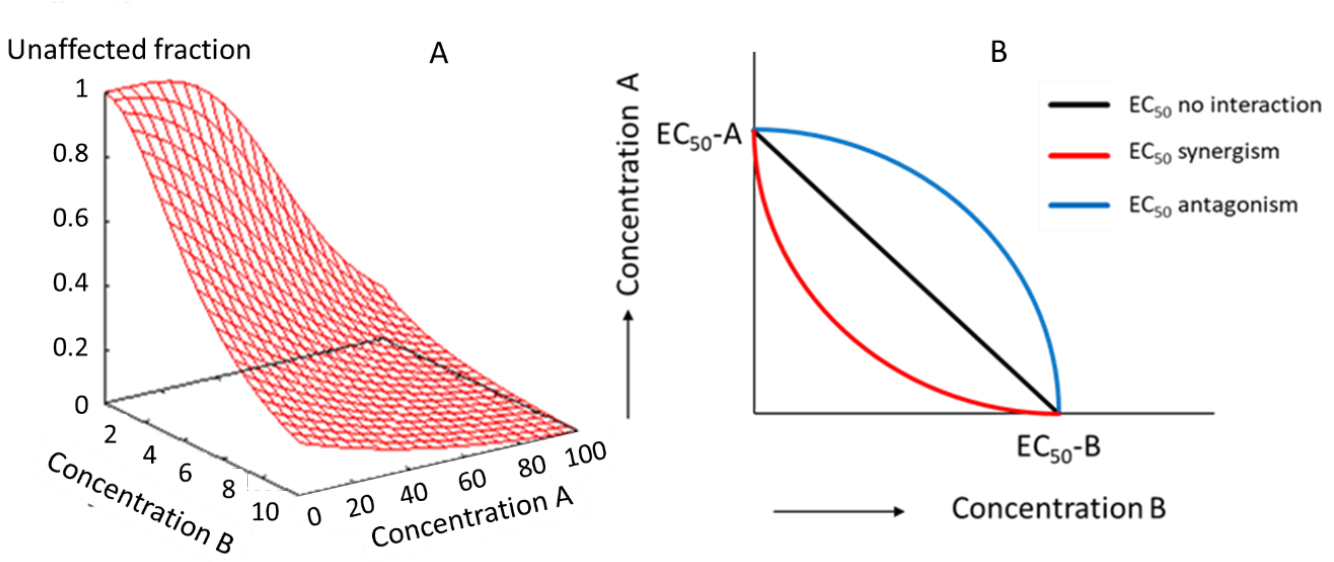


Figure 5. A. Dose-response surface showing the effect of chemicals A and B, single (sides of the surface), and in mixtures. B. Isoboles showing the toxicity of the same mixtures in a two-dimensional plane. The isoboles can be seen as a cross section through the dose-response surface. The isoboles show the 50% effect level according to Concentration Addition of mixtures of the two compounds in case they do not interact (black line), when they interact antagonistically (blue line) and when they interact synergistically (red line).

Synergism and antagonism evaluated by both concepts

The use of the terms synergism and antagonism may be problematic, because antagonism in relation to Concentration Addition (less than concentration additive; Figure 5B blue line) can simply be caused by the compounds behaving according to Response Addition, and not behaving antagonistically. Similarly, deviations from Response Addition could also mean that chemicals in the mixture do have the same mode of action, so act additively according to Concentration Addition. One can therefore only conclude on synergism/antagonism if the experimental observations are higher/lower than the predictions by *both* concepts.

Suggested further reading

Rider, C.V., Simmons, J.E. (2018). *Chemical Mixtures and Combined Chemical and Nonchemical Stressors: Exposure, Toxicity, Analysis, and Risk*, Springer International Publishing AG. ISBN-13: 978-3319562322.

Bopp, S.K., Kienzler, A., Van der Linden, S., Lamon, L., Paini, A., Parissis, N., Richarz, A.N., Triebe, J., Worth, A. (2016). Review of case studies on the human and environmental risk assessment of chemical mixtures. JRC Technical Reports EUR 27968 EN, European Union, doi:10.2788/272583.

References

Berenbaum, M.C. (1981). Criteria for analysing interactions between biologically active agents. Advances in Cancer Research 35, 269-335.

Deneer, J.W., Sinnige, T.L., Seinen, W., Hermens, J.L.M. (1988). The joint acute toxicity to *Daphnia magna* of industrial organic chemicals at low concentrations. Aquatic Toxicology 12, 33-38.

Hewlett, P.S., Plackett, R.L. (1959). A unified theory for quantal responses to mixtures of drugs: non-interactive action. Biometrics 15, 691 610.

Kraak, M.H.S., Stuijfzand, S.C., Admiraal, W. (1999). Short-term ecotoxicity of a mixture of five metals to the zebra mussel *Dreissena polymorpha*. Bulletin of Environmental Contamination and Toxicology 63, 805-812.

Van Gestel, C.A.M., Jonker, M.J., Kammenga, J.E., Laskowski, R., Svendsen, C. (Eds.) (2011). *Mixture toxicity. Linking approaches from ecological and human toxicology*. SETAC Press, Society of Environmental Toxicology and Chemistry, Pensacola.

4.4.1. Question 1

What is the motivation to perform mixture toxicity experiments?

4.4.1. Question 2

When do you expect concentration-addition and when not?

4.4.1. Question 3

What are the three possible outcomes of a mixture toxicity experiment applying concentration addition?

4.4.1. Question 4

Calculate the effect concentration at which two compounds with different modes of action showing no interaction equally contribute to a mixture causing 60% effect.

4.4.1. Question 5

One can only conclude on synergism/antagonism, if the experimental observations are higher/lower than the predictions by both concepts (concentration addition and response addition). Why?

4.4.2. Multistress Introduction

Author: Michiel Kraak

Reviewer: Kees van Gestel

Learning objectives:

You should be able to

· define stress and multistress.

· explain the ecological relevance of multistress scenarios.

Keywords: Stress, multistress, chemical-abiotic interactions, chemical-biotic interactions

Introduction

In contaminated environments, organisms are generally exposed to a wide variety of toxicants under variable and sub-optimal conditions. To gain ecological realism, multistress scenarios should thus be considered, but these are, however, understudied.

Definitions

Stress is defined as an environmental change that affects the fitness and ecological functioning of species (i.e., growth, reproduction, behaviour), ultimately leading to changes in community structure and ecosystem functioning. Multistress is subsequently defined as a situation in which an organism is exposed both to a toxicant and to stressful environmental conditions. This includes chemical-abiotic interactions, chemical-biotic interactions as well as combinations of these. Common abiotic stressors are for instance pH, drought, salinity and above all temperature, while common biotic stressors include predation, competition, population density and food shortage. Experiments on such stressors typically study, for instance, the effect of increasing temperature or the influence of food availability on the toxicity of compounds.

The present definition of multistress thus excludes mixture toxicity (see section on Mixture toxicity) as well as situations in which organisms are confronted with several suboptimal (a)biotic environmental variables jointly without being exposed to toxicants. The next chapters deal with chemical-abiotic and chemical-biotic interactions and with practical issues related with the performance of multistress experiments, respectively.

4.4.2. Question 1

Give the definitions of stress and multistress.

4.4.2. Question 2

What is the ecological relevance of testing multistress scenarios?

4.4.3. Multistress - biotic

Authors: Marjolein Van Ginneken and Lieven Bervoets

Reviewers: Michiel Kraak and Martin Holmstrup

Learning objectives:

You should be able to

- define biotic stress and to give three examples.
- explain how biotic stressors can change the toxicity of chemicals
- explain how chemicals can change the way organisms react to biotic stressors

Keywords: Multistress, chemical-biotic interactions, stressor interactions, bioavailability, behavior, energy trade-off

Introduction

Generally, organisms have to cope with the joint presence of chemical and natural stressors. Both biotic and abiotic stressors can affect the chemicals' bioavailability and toxicokinetics. Additionally, they can influence the behavior and physiology of organisms, which could result in higher or lower toxic effects. Vice versa, chemicals can alter the way organisms react to natural stressors.

By studying the effects of multiple stressors, we can identify potential synergistic, additive or antagonistic interactions, which are essential to adequately assess the risk of chemicals in nature. Relyea (2003), for instance, found that apparently safe concentrations of carbaryl can become deadly to some amphibian species when combined with predator cues. This section focuses on biotic stress, which can be defined as stress caused by living organisms and includes predation, competition, population density, food availability, pathogens and parasitism. It will describe how biotic stressors and chemicals act and interact.

Types of biotic stressors

Biotic stressors can have direct and indirect effects on organisms. For example, predators can change food web structures by consuming their prey and thus altering prey abundance and can indirectly affect prey growth and development as well, by inducing energetically-costly defense mechanisms. Also behaviors like (foraging) activity can be decreased and even morphological changes can be induced. For example, *Daphnia pulex* can develop neck spines when they are subject to predation. Similarly, parasites can alter host behavior or induce morphological changes, e.g., in coloration, but they usually do not kill their host. Yet, parasitism can compromise the immune system and alter the energy budget of the host.

High population density is a stressor that can affect energy budgets and intraspecific and interspecific competition for space, status or resources. By altering resource availability, changes in growth and size at maturity can be the result. Additionally, these competition-related stressors can affect behavior, for example by limiting the number of suitable mating partners. Also pathogens (e.g., viruses, bacteria and fungi) can lower fitness and fecundity.

It should be realized that the effects of different biotic stressors cannot be strictly separated from each other. For example, pathogens can spread more rapidly when population densities are high, while predation, on the other hand, can limit competition.

Effects of biotic stressors on bioavailability and toxicokinetics

Biotic stressors can alter the bioavailability of chemicals. For example in the aquatic environment, food level may determine the availability of chemicals to filter feeders, as they may adsorb to particulate organic matter, such as algae. As the exposure route (waterborne or via food) can influence the subsequent toxicokinetic processes, this may also change the chemicals' toxic effects.

Effects of biotic stressors on behavior

Biotic stressors have been reported to cause behavioral effects in organisms that could change the toxic effects of chemicals. These effects include altered feeding rates and reduced activities. The presence of a predator, for example, reduces prey (foraging) activity to avoid being detected by the perceived predator and so decreases chemical uptake via food. On the other hand, the condition of

the prey organisms will decrease due to the lower food consumption, which means less energy is available for other physiological processes (see below).

In addition to biotic stressors, also chemicals can disrupt essential behaviors by reduction of olfactory receptor sensitivity, cholinesterase inhibition, alterations in brain neurotransmitter levels, and impaired gonadal or thyroid hormone levels. This could lead to disruptive effects on communication, feeding rates and reproduction. An inability to find mating partners, for example, could then be worsened by a low population density. Furthermore, chemicals can alter predator-prey relationships, which might result in trophic cascades. Strong top-down effects will be observed when a predator or grazer is more sensitive to the contaminant than its prey. Alternatively, bottom-up effects are observed when the susceptibility of a prey species to predation is increased. For example, Cu exposure of fish and crustaceans can decrease their response to olfactory cues, making them unresponsive to predator stress and increasing the risk to be detected and consumed (Van Ginneken et al., 2018). Effects on the competition between species may also occur, when one species is more sensitive than the other. Thus, both chemical and biotic stressors can alter behavior and result in interactive effects that could change the entire ecosystem structure and function (Fleeger et al., 2003).

Physiology

Biotic stressors can cause elevated respiration rates of organisms, in aquatic organisms leading to a higher toxicant uptake through diffusion. On the other hand, they can also decrease respiration. For example, low food levels decrease metabolic activity and thus respiration. Additionally, a reduced metabolic rate could decrease the toxicity of chemicals which are metabolically activated. Also certain chemicals, such as metals, can cause a higher or lower oxygen consumption, which might counteract or reinforce the effects of biotic stressors.

Besides affecting respiration, both biotic and chemical stressors can induce physiological damage to organisms. For instance, predator stress and pesticides cause oxidative stress, leading to synergistic effects on the induction of antioxidant enzymes such as catalase and superoxide dismutase (Janssens and Stoks, 2013). Furthermore, the organism can eliminate or detoxify internal toxicant concentrations, e.g. by transformation via Mixed Function Oxidation enzymes (MFO) or by sequestration, i.e. binding to metallothioneins or storage in inert tissues such as granules. These defensive mechanisms for detoxification and damage control are energetically costly, leading to energy trade-offs. This means less energy can be used for other processes such as growth, locomotion or reproduction. Food availability and lipid reserves can then play an important role, as well-fed organisms that are exposed to toxicants can more easily pay the energy costs than food-deprived organisms.

Interactive effects

The possible interactions, i.e. antagonism, synergism or additivity, between effects of stressors are difficult to predict and can differ depending on the stressor combination, chemical concentration, the endpoint and the species. For *Ceriodaphnia dubia*, Qin et al. (2011) demonstrated that predator stress influenced the toxic effects of several pesticides differently. While predator cues interacted antagonistically with bifenthrin and thiacloprid, they acted synergistically with fipronil.

It should also be noted that interactive effects in nature might be weaker than when observed in the laboratory as stress levels fluctuate more rapidly or animals can move away from areas with high predator risk or chemical exposure levels. On the other hand, because generally more than two stressors are present in ecosystems, which could interact in an additive or synergistic way as well, they might be even more important in nature. Understanding interactions among multiple stressors is thus essential to estimate the actual impact of chemicals in nature.

References

Fleeger, J.W., Carman, K.R., Nisbet, R.M. (2003). Indirect effects of contaminants in aquatic ecosystems. Science of the Total Environment 317, 207-233.

Janssens, L., Stoks, R. (2013). Synergistic effects between pesticide stress and predator cues: conflicting results from life history and physiology in the damselfly *Enallagma cyathigerum*. Aquatic Toxicology 132, 92-99.

Qin, G., Presley, S.M., Anderson, T.A., Gao, W., Maul, J.D. (2011). Effects of predator cues on pesticide toxicity: toward an understanding of the mechanism of the interaction. Environmental Toxicology and Chemistry 30, 1926-1934.

Relyea, R.A. (2003). Predator cues and pesticides: a double dose of danger for amphibians. Ecological Applications 13, 1515-1521.

Van Ginneken, M., Blust, R., Bervoets, L. (2018). Combined effects of metal mixtures and predator stress on the freshwater isopod *Asellus aquaticus*. Aquatic Toxicology 200, 148-157.

4.4.3. Question 1

Give the definition of biotic stress and give 3 examples.

4.4.3. Question 2

How can biotic stressors change the toxic effects of chemicals?

4.4.3. Question 3

Give an example of how chemicals can change the way organisms react to biotic stress?

4.4.4. Multistress - abiotic

Author: Martina Vijver

Reviewers: Kees van Gestel, Michiel Kraak, Martin Holmstrup

Learning objectives:

You should be able to

- relate stress to the ecological niche concept
- list abiotic factors that may alter the toxic effects of chemicals on organisms, and indicate if these abiotic factors decrease or increase the toxic effects of chemicals on organisms

Key words: Stress, ecological niche concept, multistress, chemical-abiotic interactions

Introduction: stress related to the ecological niche concept

The concept of stress can be defined at various levels of biological organization, from biochemistry to species fitness, ultimately leading to changes in community structure and ecosystem functioning. Yet, stress is most often studied in the context of individual organisms. The concept of stress is not absolute and can only be defined with reference to the normal range of ecological functioning. This is the case when organisms are within their range of tolerance (so-called ecological amplitude) or within their ecological niche, which describes the match of a species to specific environmental conditions. Applying this concept to stress allows it to be defined as a condition evoked in an organism by one or more environmental factors that bring the organism near or over the edges of its ecological niche (Van Straalen, 2003), see Figure 1.

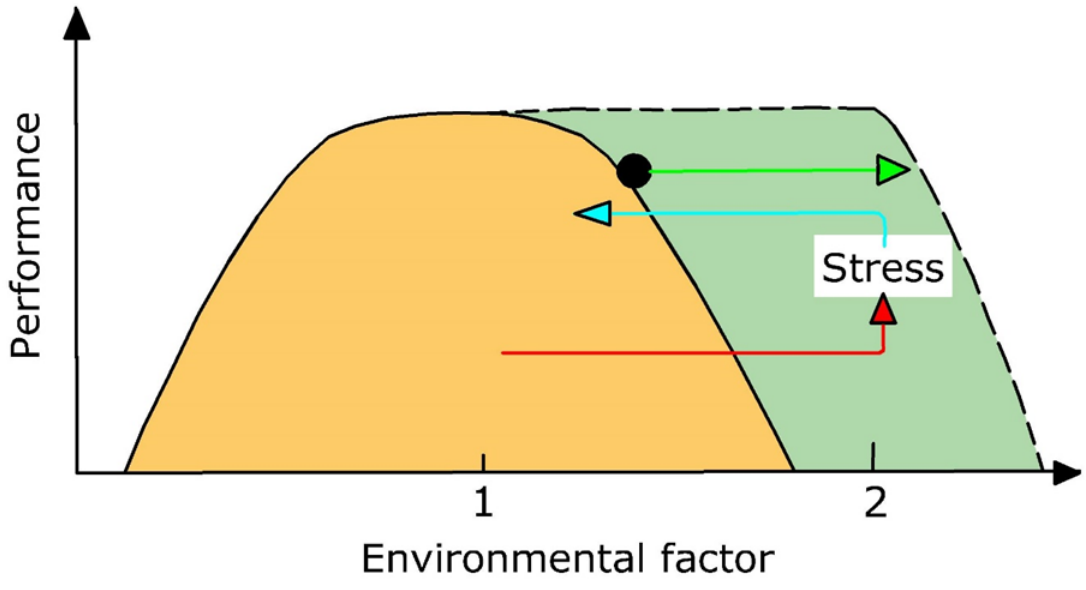


Figure 1. Schematic illustration of the ecological amplitude or niche-based (bell shaped curve) definition of stress. Stress arises when an environmental

factor increases from point 1 to point 2 (red line) and the species is forced outside its ecological niche. By definition, the organism cannot grow and reproduce outside this niche, but it may survive there temporarily, if it can return in time to its niche (blue line). If the borders of the niche are extended through adaptation, then this specific state of the environmental factor does not result in stress anymore and the performance of the species falls within the normal operating range at that condition (green line). Redrawn from Van Straalen (2003) by Wilma IJzerman.

Multistress is subsequently defined as a situation in which an organism is exposed both to a toxicant and to stressful environmental conditions (see section Multistress - Introduction and definitions). This includes chemical-abiotic interactions, chemical-biotic interactions (see section Multistress - chemical - biotic interactions) as well as combinations of these. In general, organisms living under conditions close to their environmental tolerance limits appear to be more vulnerable to additional chemical stress. The opposite also holds: if organisms are stressed due to exposure to elevated levels of contaminants, their ability to cope with sub-optimal environmental conditions is reduced.

Chemical-abiotic interactions

Temperature. One of the predominant environmental factors altering toxic effects obviously is temperature. For poikilothermic (cold-blooded) organisms, increases in temperature lead to an increase in activity, which may affect both the uptake and the effects of chemicals. In a review by Heugens et al. (2001), studies reporting the effect chemicals on aquatic organisms in combination with abiotic factors like temperature, nutritional state and salinity were discussed. Generally, toxic effects increased with increasing temperature. Dependent on the effect parameter studied, the differences in toxic effects between laboratory and relevant field temperatures ranged from a factor of 2 to 130.

Also freezing temperatures may interfere with chemical effects as was shown in another influential review of Holmstrup et al. (2010). Membrane damage is mentioned as an explanation for the synergistic interaction between combinations of metals and temperatures below zero.

Food. Food availability may have a strong effect on the sensitivity of organisms to chemicals (see section Multistress - chemical - biotic interactions). In general decreasing food or nutrient levels increased toxicity, resulting in differences in toxicity between laboratory and relevant field situations ranging from a factor of 1.2 to 10 (Heugens et al., 2001). Yet, way higher differences in toxic effects related to food levels have been reported as well: Experiments performed with daphnids in cages that were placed in outdoor mesocosm ditches (see sections on Cosm studies and In situ bioassays) showed stunning differences in sensitivity to the insecticide thiacloprid. Under conditions of low to ambient nutrient concentrations, the observed toxicity, expressed as the lowest observed effect concentration (LOEC) for growth and reproduction occurred at thiacloprid concentrations that were 2500-fold lower than laboratory-derived LOEC values. Contrary to the low nutrient treatment, such altered toxicity was often not observed under nutrient-enriched conditions (Barmentlo et al submitted). The difference was likely attributable to the increased primary production that allowed for compensatory feeding and perhaps also reduced the bioavailability of the insecticide. Similar results were observed for sub-lethal endpoints measured on the damselfly species *Ischnura elegans*, for which the response to thiacloprid exposure strongly depended on food availability and quality. Damselfies that were feeding on natural resources were significantly more affected than those that were offered high quality artificial food (Barmentlo et al submitted).

Salinity. The influence of salinity on toxicity is less clear (Heugens et al. 2001). If salinity pushes the organism towards its niche boundaries, it will worsen the toxic effects that it is experiencing. In case that a specific salinity fits in the ecological niche of the organism, processes affecting exposure will predominantly determine the stress it will experience. This for instance means that metal toxicity decreases with increasing salinity, as it is strongly affected by the competition of ions (see section on Metal speciation). The toxic effect induced by organophosphate insecticides however, increases with increasing salinity. For other chemicals, no clear relationship between toxicity and salinity was observed. A salinity increase from freshwater to marine water decreased toxicity by a factor of 2.1 (Heugens et al. 2001). However, as less extreme salinity changes are more relevant under field conditions, the change in toxicity is probably much smaller.

pH. Many organisms have a species-specific range of pH levels at which they function optimally. At pH values outside the optimal range, organisms may show reduced reproduction and growth, in extreme cases even reduced survival. In some cases, the effects of pH may be indirect, as pH may also have an important impact on exposure of organisms to toxicants. This is especially the case for metals and ionizable chemicals: metal speciation, but also the form in which ionizable chemicals occur in the environment and therefore their bioavailability, is highly dependent on pH (see sections on Metal speciation and Ionogenic organic chemicals). An example of the interaction between pH and metal effects was shown by Crommentuijn et al. (1997), who observed a reduced

control reproduction of the springtail *Folsomia candida*, but also the lowest cadmium toxicity at a soil pH_{KCl} 7.0 compared to pH_{KCl} 3.1-5.7.

Drought. In soil, the moisture content (see section on Soil) is an important factor, since drought is often limiting the suitability of the soil as a habitat for organisms. Holmstrup et al. (2010), reviewing the literature, concluded that chemicals interfering with the drought tolerance of soil organisms, e.g. by affecting the functioning of membranes or the accumulation of sugars, may exacerbate the effects of drought. Earthworms are breathing through the skin and can only survive in moist soils, and the eggs of springtails can only survive at a relative air humidity close to 100%. This makes these organisms especially sensitive to drought, which may be enhanced by exposure to chemicals like metals, polycyclic aromatic hydrocarbon or surfactants (Holmstrup et al., 2010).

Many different abiotic conditions, such as oxygen levels, light, turbidity, and organic matter content, can push organisms towards the boundaries of their niche, but we will not discuss all stressors in this book.

Multistress in environmental risk assessment

In environmental risk assessment, differences between stress-induced effects as determined in the laboratory under standardized optimal conditions with a single toxicant and the effects induced by multiple stressors are taken into account by applying an uncertainty factor. Yet, the choice for uncertainty factors is based on little ecological evidence. In 2001, Heugens already argued for obtaining uncertainty factors that sufficiently protect natural systems without being overprotective. Van Straalen (2003) echoed this and in current research the question is still raised if enough understanding has been gained to make accurate laboratory-to-field extrapolations. It remains a challenge to predict toxicant-induced effects on species and even on communities while accounting for variable and suboptimal environmental conditions, even though these conditions are common aspects of natural ecosystems (see for instance the section on Eco-epidemiology).

References

Barmentlo, S.H., Vriend, L.M, van Grunsven, R.H.A., Vijver, M.G. (submitted). Evidence that neonicotinoids contribute to damselfly decline.

Crommentuijn, T., Doornekamp, A., Van Gestel, C.A.M. (1997). Bioavailability and ecological effects of cadmium on *Folsomia candida* (Willem) in an artificial soil substrate as influenced by pH and organic matter. Applied Soil Ecology 5, 261-271.

Heugens, E.H., Hendriks, A.J., Dekker, T., Van Straalen, N.M., Admiraal, W. (2001). A review of the effects of multiple stressors on aquatic organisms and analysis of uncertainty factors for use in risk assessment. Critical Reviews in Toxicology 31, 247-84.

Holmstrup, M., Bindesbøl, A.M., Oostingh, G.J., Duschl, A., Scheil, V., Köhler, H.R., Loureiro, S., Soares, A.M.V.M., Ferreira, A.L.G., Kienle, C., Gerhardt, A., Laskowski, R., Kramarz, P.E., Bayley, M., Svendsen, C., Spurgeon, D.J. (2010). Review Interactions between effects of environmental chemicals and natural stressors: A review. Science of the Total Environment 408, 3746-3762.

Van Straalen, N.M. (2003). Ecotoxicology becomes stress ecology Environmental Science and Technology 37, 324A-330A.

4.4.4. Question 1

1. Describe the niche-based definition of stress using Figure 1.

a. Explain what happens to a species when it has to deal with a temporary stress

b. Explain what happens to a species if the stress is long term and the species is able to adapt to it

---

4.4.4. Question 2

Mention different abiotic factors and indicate how they may affect the sensitivity of organisms to chemicals.

4.4.5. Chronic toxicity - Eco

Author: Michiel Kraak

Reviewers: Kees van Gestel and Lieven Bervoets

Learning objectives:

You should be able to

- explain the concepts involved in chronic toxicity testing, including the Acute to Chronic Ratio (ACR).
- design chronic toxicity experiments and to solve the challenges involved in chronic toxicity testing.
- interpret the results of chronic toxicity experiments and to mention the types of effects of toxicants that cannot be determined in acute toxicity experiments.

Key words: Chronic toxicity, chronic sublethal endpoints, Acute to Chronic Ratio, mode of action.

Introduction

Most toxicity tests performed are short-term high-dose experiments, acute tests in which mortality is often the only endpoint. This is in sharp contrast with the field situation, where organisms are often exposed to relatively low levels of contaminants for their entire life span. The shorter the life cycle of the organism, the more realistic this scenario becomes. Hence, there is an urgent need for chronic toxicity testing. It should be realized though, that the terms acute and chronic have to be considered in relation to the length of the life cycle of the organism. A short-term exposure of four days is acute for fish, but chronic for algae, comprising already four generations.

From acute to chronic toxicity testing

The reason for the bias towards acute toxicity testing is obviously the higher costs involved in chronic toxicity testing, simply caused by the much longer duration of the test. Yet, chronic toxicity testing is challenging for several other reasons as well. First of all, during prolonged exposure organisms have to be fed. Although unavoidable, especially in aquatic toxicity testing, this will definitely influence the partitioning and the bioavailability of the test compound. Especially lipophilic compounds will strongly bind to the food, making toxicant uptake via the food more important than for hydrophilic compounds, thus causing compound specific changes in exposure routes. For chronic aquatic toxicity tests, especially for sediment testing, it may be challenging to maintain sufficiently high oxygen concentrations throughout the entire experiment (Figure 1).

Figure 1. Experimental design of a chronic sediment toxicity experiment, showing the experimental units and the aeration system.

Obvious choices to be made include the duration of the exposure and the endpoints of the test. Generally it is aimed at including at least one reproductive event or the completion of an entire life cycle of the organism within the test duration. To ensure this, validity criteria are set to the different test guidelines, such as:

- the mean number of living offspring produced per control parent daphnid surviving till the end of the test should be above 60 (OECD, 2012).

- 85% of the adult control chironomid midges from the control treatment should emerge between 12 and 23 days after the start of the experiment (OECD, 2010).

- the mean number of juveniles produced by 10 control collembolans should be at least 100 (OECD, 2016a).

Chronic toxicity

Generally toxicity increases with increasing exposure time, often expressed as the acute-to-chronic ratio (ACR), which is defined as the LC_{50} from an acute test divided by the door NOEC of EC_{10} from the chronic test. Alternatively, as shown in Figure 2, the acute LC_{50} can be divided by the chronic LC_{50}. If compounds exhibit a strong direct lethal effect, the ACR will be low, but for compounds that slowly build up lethal body burdens (see section on Critical body concentrations) it can be very high. Hence, there is a relationship between the mode of action of a compound and the ACR. Yet, if chronic toxicity has to be extrapolated from acute toxicity data and the mode of action of the compound is unknown, an ACR of 10 is generally considered. It should be realized though that this number is chosen quite arbitrarily, potentially leading to under- as well as over estimation of the actual ACR.

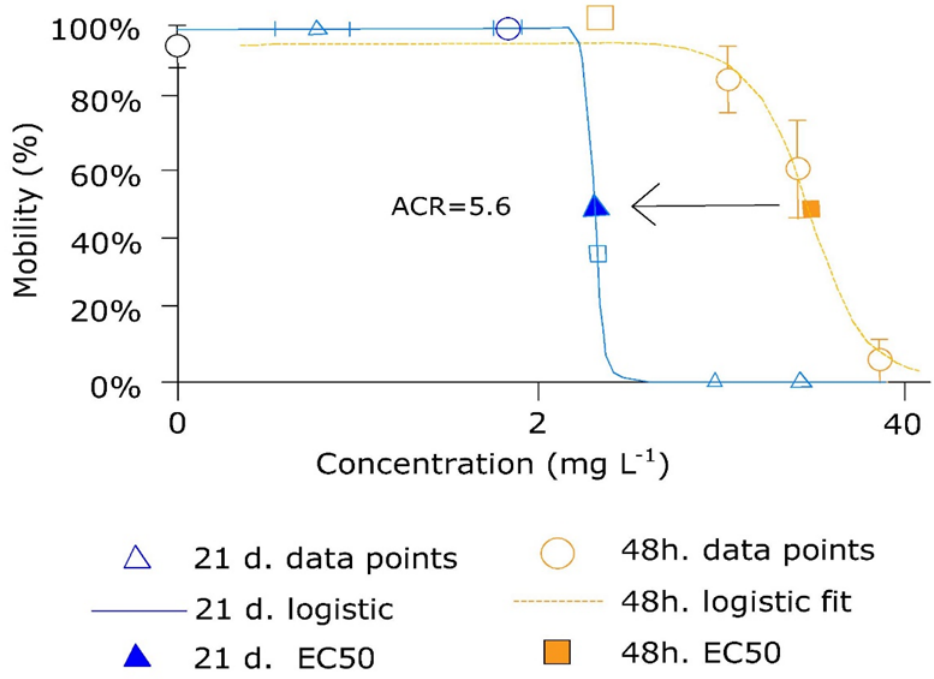


Figure 2. Average mobility (% of initial animals) of Daphnia magna (n = 15) exposed to a concentration range of the flame retardant ALPI (mg L^{-1}) in Elendt medium after 48 h (± s.e. in x and y, n = 4 × 5 individuals per concentration) and after 21 days (± s.e. in x, n = 15 individuals per concentration). The toxicity increases with increasing exposure time with an Acute Chronic Ratio (ACR) of 5.6. Redrawn from Waaijers at al. (2013) by Wilma IJzerman.

Since reproductive events and the completion of life cycles are involved, chronic toxicity tests allow an array of sublethal endpoints to be assessed, including growth and reproduction, as well as species specific endpoints like emergence (time) of chironomids. Consequently, compounds with different modes of action may cause very diverse sublethal effects on the test organisms during chronic exposure (Figure 3). The polycyclic aromatic compound (PAC) phenanthrene did not affect the completion of the life cycle of the midges, but above a certain exposure concentration the larvae died and no emergence was observed at all, suggesting a non-specific mode of action (narcosis). In contrast, the PAC acridone caused no mortality but delayed adult emergence significantly over a wide range of test concentrations, suggesting a specific mode of action affecting life cycle parameters of the midges (Leon Paumen et al., 2008). This clearly demonstrates that specific effects on life cycle parameters of compounds with different modes of action need time to become expressed.

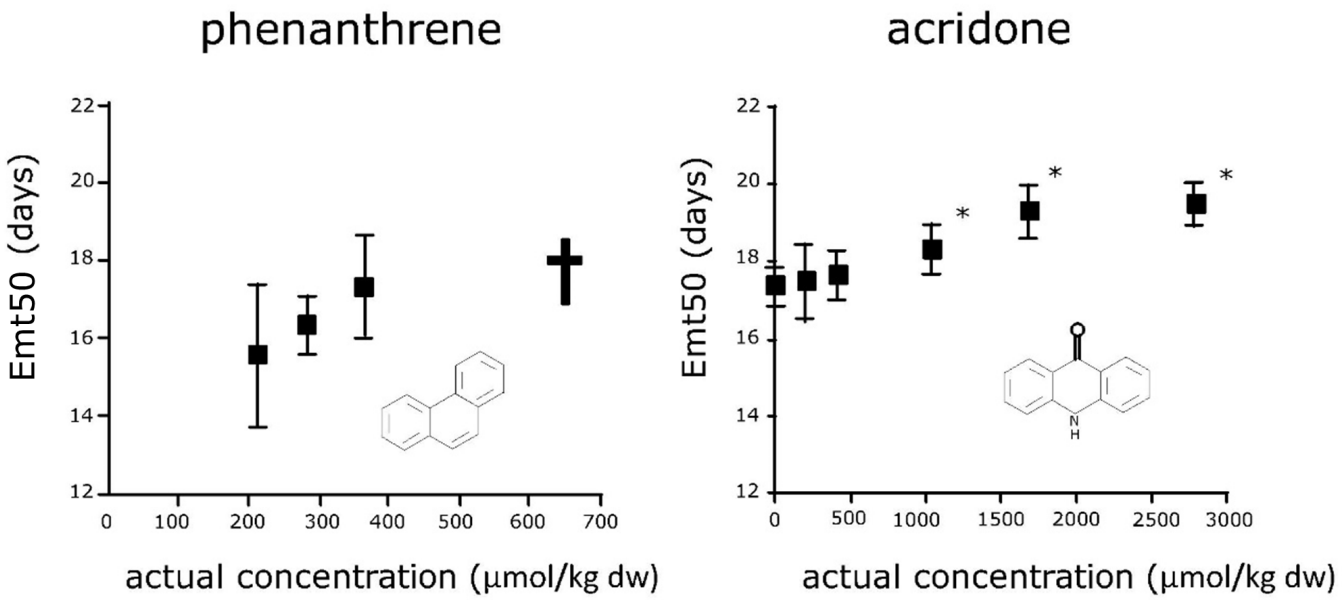


Figure 3. Effect of two polycyclic aromatic compounds (PACs) on the emergence time of Chironomus riparius males from spiked sediments. X-axis: actual concentrations of the compounds measured in the sediment. Y-axis: 50% male emergence time (EMt_{50}, days, average plus standard deviations). †Concentrations with no emerging midges. *EMt_{50} value significantly different from control value ($p < 0.05$). Redrawn from Leon Paumen et al. (2008) by Wilma IJzerman.

Chronic toxicity tests are single species tests, but if the effects of toxicants are assessed on all relevant life-cycle parameters, these can be integrated into effects on population growth rate (*r*). For the 21-day daphnid test this is achieved by the integration of age-specific data on the probability of survival and fecundity. The population growth rates calculated from chronic toxicity data are obviously not related to natural population growth rates in the field, but they do allow to construct dose-response relationships for the effects of toxicants on *r*, the ultimate endpoint in chronic toxicity testing (Figure 4; Waaijers et al., 2013).

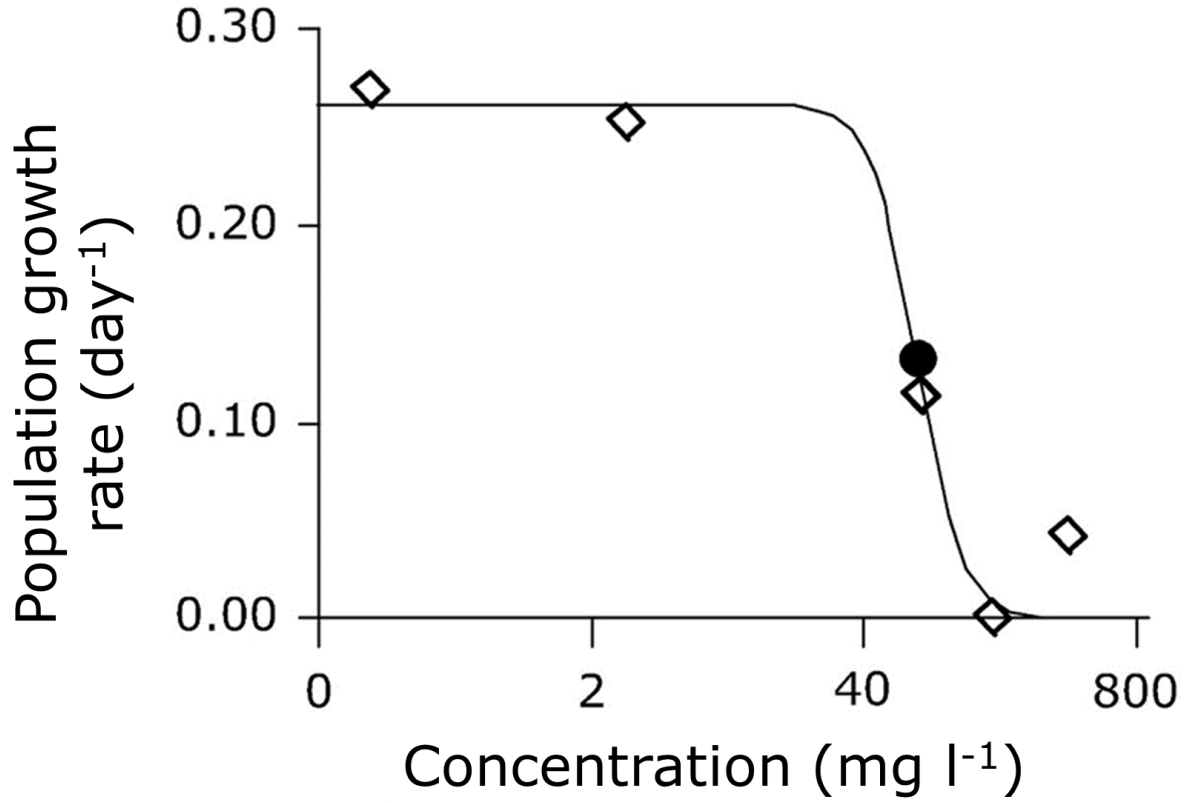


Figure 4: Population growth rate (d^{-1}) of Daphnia magna (n = 15) exposed to a concentration range of the flame retardant DOPO (mg L^{-1}) in Elendt medium after 21 days. The average population growth rate rate (◊) is shown (s.e. in x and y are smaller than the data points and therefore omitted). The EC_{50} is plotted as ● (s.e. smaller than data point) and the logistic curve represents the fitted concentration−response relationship. Redrawn from Waaijers et al. (2013) by Wilma IJzerman.

Chronic toxicity testing in practice

Several protocols for standardized chronic toxicity tests are available, although less numerous than for acute toxicity testing. For water, the most common test is the 21 day *Daphnia* reproduction test (OECD, 2012), for sediment 28-day test guidelines are available for the midge *Chironomus riparius* (OECD, 2010) and for the worm *Lumbriculus variegatus* (OECD, 2007). For terrestrial soil, the springtail *Folsomia candida* (OECD, 2016a) and the earthworm *Eisenia fetida* (OECD, 2016b) are the most common test species, but also for enchytraeids a reproduction toxicity test guideline is available (OECD, 2016c). For a complete overview see (https://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-2-effects-on-biotic-systems_20745761/datedesc#collectionsort).

References

Leon Paumen, M., Borgman, E., Kraak, M.H.S., Van Gestel, C.A.M., Admiraal, W. (2008). Life cycle responses of the midge *Chironomus riparius* to polycyclic aromatic compound exposure. Environmental Pollution 152, 225-232.

OECD (2007). OECD Guideline for Testing of Chemicals. Test No. 225: Sediment-Water *Lumbriculus* Toxicity Test Using Spiked Sediment. Section 2: Effects on Biotic Systems; Organization for Economic Co-operation and Development: Paris, 2007.

OECD (2010). OECD Guideline for Testing of Chemicals. Test No. 233: Sediment-Water Chironomid Life-Cycle Toxicity Test Using Spiked Water or Spiked Sediment. Section 2: Effects on Biotic Systems; Organization for Economic Co-operation and Development: Paris, 2010.

OECD (2012). OECD Guideline for Testing of Chemicals. Test No. 211. *Daphnia magna* Reproduction Test No. 211. Section 2: Effects on Biotic Systems; Organization for Economic Co-operation and Development: Paris, 2012.

OECD (2016a). OECD Guideline for Testing of Chemicals. Test No. 232. Collembolan Reproduction Test in Soil. Section 2: Effects on Biotic Systems; Organization for Economic Co-operation and Development: Paris, 2016.

OECD (2016b). OECD Guideline for Testing of Chemicals. Test No. 222. Earthworm Reproduction Test (*Eisenia fetida/Eisenia andrei*). Section 2: Effects on Biotic Systems; Organization for Economic Co-operation and Development: Paris, 2016.

OECD (2016c). OECD Guideline for Testing of Chemicals. Test No. 220: Enchytraeid Reproduction Test. Section 2: Effects on Biotic Systems; Organization for Economic Co-operation and Development: Paris, 2016.

Waaijers, S.L., Bleyenberg, T.E., Dits, A., Schoorl, M., Schütt, J., Kools, S.A.E., De Voogt, P., Admiraal, W., Parsons, J.R., Kraak, M.H.S. (2013). Daphnid life cycle responses to new generation flame retardants. Environmental Science and Technology 47, 13798-13803.

4.4.5. Question 1

In acute toxicity tests the LC50 is derived after a short exposure time. Mention two outcomes of chronic toxicity tests that cannot be determined in acute toxicity tests.

4.4.5. Question 2

For which mode of toxic action of compounds do you expect low Acute-to-Chronic Ratios (ACR) and for which chemical mode of action do you expect ACR to be high?

4.4.6. Multigeneration toxicity testing - Eco

Author: Michiel Kraak

Reviewers: Kees van Gestel, Miriam Leon Paumen

Learning objectives:

You should be able to

· explain how effects of toxicants may propagate during multigeneration exposure.

· describe the experimental challenges and limitations of multigeneration toxicity testing and to be able to design multigeneration tests.

· explain the implications of multigeneration testing for ecological risk assessment.

Key words: Multigeneration exposure, extinction, adaptation, test design

Introduction

It is generally assumed that chronic life cycle toxicity tests are indicative of the actual risk that populations suffer from long-term exposure. Yet, at contaminated sites, organisms may be exposed during multiple generations and the shorter the life cycle of the organism, the more realistic this scenario becomes. There are, however, only few multigeneration studies performed, due to the obvious time and cost constraints. Since both aquatic and terrestrial life cycle toxicity tests generally last for 28 days (see section on Chronic toxicity), multigeneration testing will take approximately one month per generation. Moreover, the test compound often affects the life cycle of the test species in a dose-dependent manner. Consequently, the control population, for example, could already be in the 9^{th} generation, while an exposed population could still be in the 8^{th} generation due to chemical exposure related delay in growth and/or development. On top of these experimental challenges, multigeneration experiments are extremely error prone, simply because the chance that an experiment fails increases with increasing exposure time.

Experimental considerations

Designing a multigeneration toxicity experiment is challenging. First of all, there is the choice of how many generations the experiment should last, which is most frequently, but completely arbitrarily, set at approximately 10. Test concentrations have to be chosen as well, mostly based on chronic life cycle EC_{50} and EC_{10} values (Leon Paumen et al. 2008). Yet, it cannot be anticipated if, and to what extent, toxicity increases (or decreases) during multigeneration exposure. Hence, testing only one or two exposure concentrations increases the risk that the observed effects are not dose related, but are simply due to stochasticity. If the test

concentrations chosen are too high, many treatments may go extinct after few generations. In contrast, too low test concentrations may show no effect at all. The latter was observed by Marinkovic at al. (2012), who had to increase the exposure concentrations during the experiment (see Figure in graphical abstract of Marinkovic et al., 2012). Finally, since a single experimental treatment often consists of an entire population, treatment replication is also challenging.

Once the experiment is running, choices have to be made on the transition from generation to generation. If a replicate is maintained in a single jar, vessel or aquarium, generations may overlap and exposure concentrations may decrease with time. Therefore, most often a new generation is started by exposing offspring from the previous exposed parental generation in a freshly spiked experimental unit.

If the aim is to determine how a population recovers when the concentration of the toxicant decreases with time, exposure to a single spiked medium also is an option, which seems most applicable to soils (Ernst et al., 2016; van Gestel et al., 2017). To assess recovery after several generations of (continuous) exposure to contaminated media, offspring from previous exposed generations may be maintained under control conditions.

A wide variety of endpoints can be selected in multigeneration experiments. In case of aquatic insects like the non-biting midge *Chironomus riparius* these include survival, larval development time, emergence, emergence time, adult life span and reproduction. For terrestrial invertebrates survival, growth and reproduction can be selected. Only a very limited number of studies evaluated actual population endpoints like population growth rate (Postma and Davids, 1995).

To persist or to perish

If organisms are exposed for multiple generations the effects tend to worsen, ultimately leading to extinction, first of the population exposed to the highest concentration, followed by populations exposed to lower concentrations in later generations (Leon Paumen et al. 2008). Yet, it cannot be excluded that extinction occurs due to the relatively small population sizes in multigeneration experiments, while larger populations may pass a bottleneck and recover during later generations.

Thresholds have also been reported, as shown in Figure 1 (Leon Paumen et al. 2008). Below certain exposure concentrations the exposed populations perform equally well as the controls, generation after generation. Hence, these concentrations may be considered as the 'infinite no effect concentration'. A mechanistic explanation may be that the metabolic machinery of the organism is capable of detoxifying or excreting the toxicants and that this takes so little energy that there is no trade off regarding growth and reproduction.

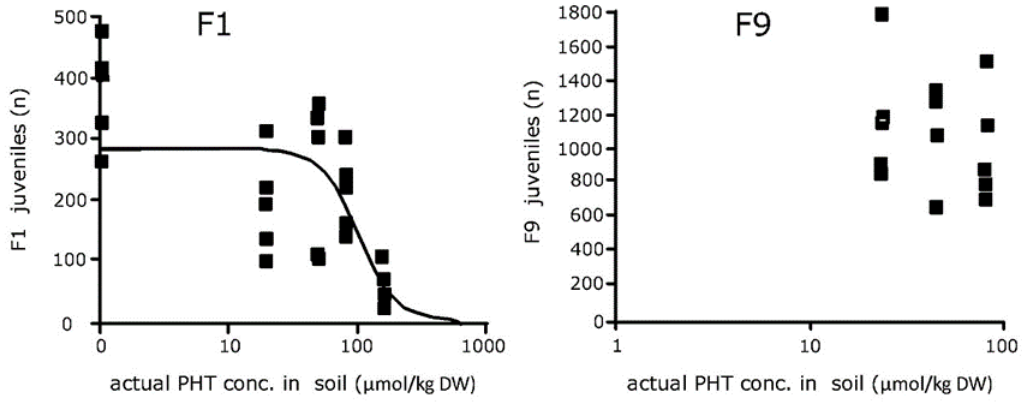


Figure 1. Transition from dose-response relationships to threshold concentrations during a multigeneration toxicity experiment with the collembolan Folsomia candida. Redrawn from Leon Paumen et al. (2008) by Wilma IJzerman.

It is concluded that the frequently reported worsening of effects during multigeneration toxicant exposure raises concerns about the use of single-generation studies in risk assessment to tackle long-term population effects of environmental toxicants.

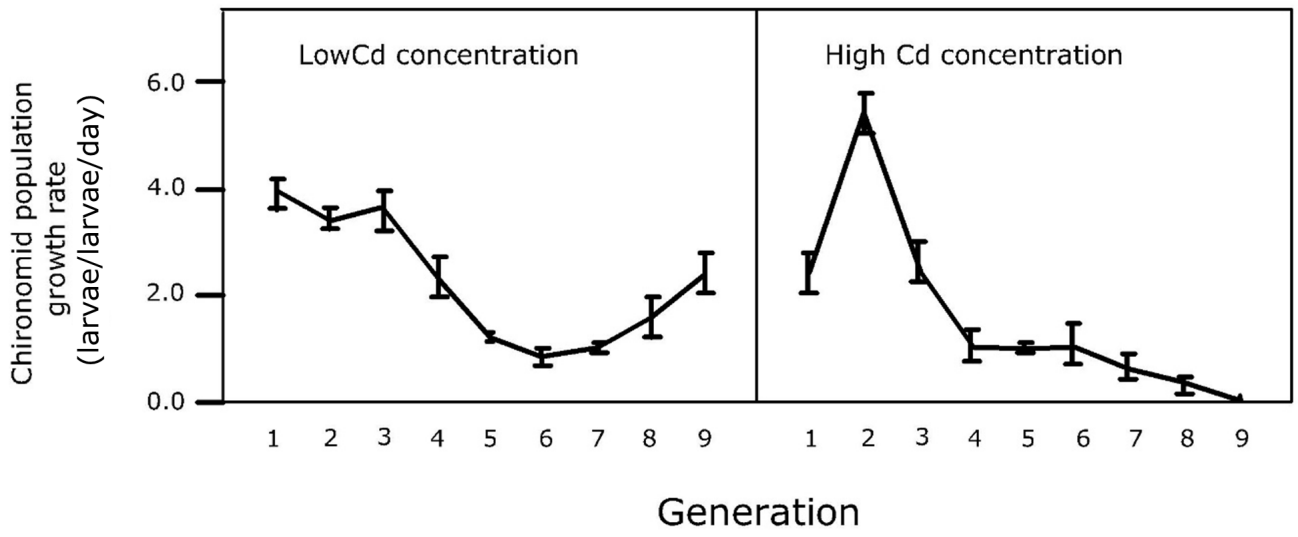


Figure 2. Extinction at a relatively high exposure concentration and adaptation at a relatively low exposure concentration during a multigeneration toxicity experiment with the non-biting midge Chironomus riparius. Redrawn from Postma & Davids (1995) by Wilma IJzerman.

If populations exposed for multiple generations do not get extinct and persist, they may have developed resistance or adaptation (Figure 2). Regular sensitivity testing can therefore be included in multigeneration experiments, as depicted in Figure 1. Yet, it is still under debate whether this lower sensitivity is due to genetic adaptation, epigenetics or phenotypic plasticity (Marinkovic et al., 2012).

References

Ernst, G., Kabouw, P., Barth, M., Marx, M.T., Frommholz, U., Royer, S., Friedrich, S. (2016). Assessing the potential for intrinsic recovery in a Collembola two-generation study: possible implementation in a tiered soil risk assessment approach for plant protection products. Ecotoxicology 25, 1-14.

Leon Paumen, M., Steenbergen, E., Kraak, M.H.S., Van Straalen, N. M., Van Gestel, C.A.M. (2008). Multigeneration exposure of the springtail *Folsomia candida* to phenanthrene: from dose-response relationships to threshold concentrations. Environmental Science and Technology 42, 6985-6990.

Marinkovic, M., De Bruijn, K., Asselman, M., Bogaert, M., Jonker, M.J., Kraak, M.H.S., Admiraal, W. (2012). Response of the nonbiting midge *Chironomus riparius* to multigeneration toxicant exposure. Environmental Science and Technology 46, 12105−12111.

Postma. J.F., Davids, C. (1995). Tolerance induction and life-cycle changes in cadmium-exposed *Chironomus riparius* (Diptera) during consecutive generations. Ecotoxicology and Environmental Safety 30, 195-202.

Van Gestel, C.A.M., De Lima e Silva, C., Lam, T., Koekkoek, J.C. Lamoree, M.H., Verwei, R.A. (2017). Multigeneration toxicity of imidacloprid and thiacloprid to *Folsomia candida*. Ecotoxicology 26, 320-328.

4.4.6. Question 1

What is the motivation to perform multigeneration experiments?

4.4.6. Question 2

What are the two alternative outcomes of multigeneration toxicity experiments?

4.4.6. Question 3

What are the implications of multigeneration testing for ecological risk assessment?

4.4.7. Tropical Ecotoxicology

Authors: Michiel Daam, Jörg Römbke

Reviewer: Kees van Gestel, Michiel Kraak

Learning objectives:

You should be able

· to name the distinctive features of tropical and temperate ecosystems

· to explain their implications for environmental risk assessment in these regions

· to mention some of the main research needs in tropical ecotoxicology

Key words: Environmental risk assessment; pesticides; temperature; contaminant fate; test methods

Introduction

The tropics cover the area of the world (approx. 40%) that lies between the Tropic of Cancer, 23½° north of the equator and the Tropic of Capricorn, 23½° south of the equator. It is characterized by, on average, higher temperatures and sunlight levels than in temperate regions. Based on precipitation patterns, three main tropical climates may be distinguished: Tropical rainforest, monsoon and savanna climates. Due to the intrinsic differences between tropical and temperate regions, differences in the risks of chemicals are also likely to occur. These differences are briefly exemplified by taking pesticides as an example, addressing the following subjects: 1) Climate-related factors; 2) Species sensitivities; 3) Testing methods; 4) Agricultural practices and legislation.

1. Climate-related factors

Three basic climate factors are essential for pesticide risks when comparing temperate and tropical aquatic agroecosystems: rainfall, temperature and sunlight. For example, high tropical temperatures have been associated with higher microbial activities and hence enhanced microbial pesticide degradation, resulting in lower exposure levels. On the other hand, toxicity of pesticides to aquatic biota may be higher with increasing temperature. Regarding terrestrial ecosystems, other important abiotic factors to be considered are soil humidity, pH, clay and organic carbon content and ion exchange capacity (i.e. the capacity of a soil to adsorb certain compounds) (Daam et al., 2019). Although several differences in climatic factors may be distinguished between tropical and temperate areas, these do not lead to consistent greater or lesser pesticide risk (e.g. Figure 1).

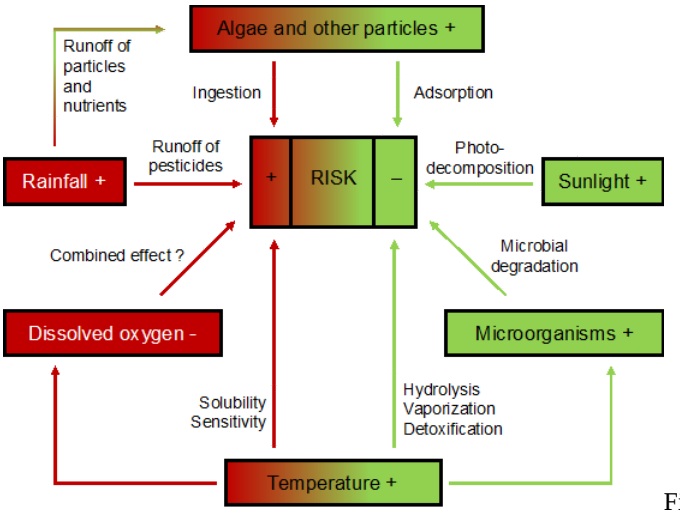


Figure 1. Schematic overview of the climatic related factors that have a possible influence on the risks of pesticides to aquatic ecosystems. The "+" and "-" in the parameter textboxes indicate relatively higher and lower levels of these parameters in tropical compared to temperate regions, respectively. Similarly, the "+" and "-" in the textbox "RISK" indicate a higher and lower risk in tropical compared to temperate freshwaters, respectively. Adapted from Daam and Van den Brink (2010).

2. Species sensitivities

Tropical areas harbour the highest biodiversity in the world and generate nearly 60% of the primary production. This higher species richness, as compared to their temperate counterparts, dictates that the possible occurrence of more sensitive species cannot be ignored. However, studies comparing the sensitivity of species from the same taxonomic group did not demonstrate a consistent higher or lower sensitivity of tropical organisms compared to temperate organisms (e.g. Figure 2).

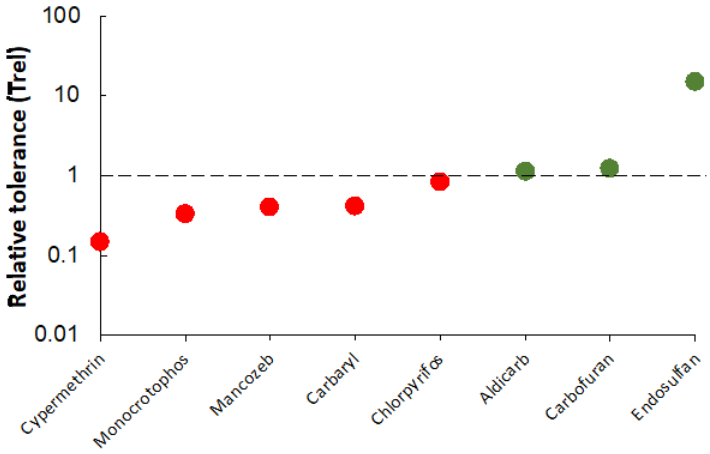


Figure 2. Comparison of the pesticide sensitivity of the tropical earthworm Perionyx excavatus with that of Eisenia fetida sensu lato using the relative tolerance (Trel) approach. The vertical dashed line at Trel = 1 indicates the sensitivity of E. fetida sensu lato. A Trel < 1 (red dots) and Trel > 1 (green dots) indicate a higher and lower sensitivity of *P. excavatus* relative to *E. fetida* sensu lato, respectively. PAF = potentially affected fraction. Modified from Daam et al. (2019).

3) Testing methods

Given the vast differences in environmental conditions between tropical and temperate regions, the use of test procedures developed under temperate environments to assess pesticide risks in tropical areas has often been disputed. Subsequently, methods developed under temperate conditions need to be adapted to tropical environmental conditions, e.g. by using tropical test substrates and by testing at higher temperatures (Niva et al., 2016). As discussed above, tropical and temperate species from the same taxonomic group are not expected to demonstrate consistent differences in sensitivity. However, certain taxonomic groups may be more represented and/or ecologically or economically more important in tropical areas, such as freshwater shrimps (Daam and Rico, 2016) and (terrestrial) termites (Daam et al., 2019). Subsequently, the development of test procedures for such species and the incorporation in risk assessment procedures seems imperative.

4) Agricultural practices and legislation

Agricultural practices in tropical countries are likely to lead to a higher pesticide exposure and hence higher risks to aquatic and terrestrial ecosystems under tropical conditions. Some of the main reasons for this include i) unnecessary applications and overuse; ii) use of cheaper but more hazardous pesticides, and iii) dangerous transportation and storage conditions, all often a result of a lack in training of pesticide applicators in the tropics (Daam and Van den Brink, 2010; Daam et al., 2019). Finally, countries in tropical regions usually do not have strict laws and risk assessment regulations in place regarding the registration and use of pesticides, meaning that pesticides banned in temperate regions for environmental reasons are often regularly available and used in tropical countries such as Brazil (e.g. Waichman et al. 2002).

References and recommended further reading

Daam, M.A., Van den Brink, P.J. (2010). Implications of differences between temperate and tropical freshwater ecosystems for the ecological risk assessment of pesticides. Ecotoxicology 19, 24-37.

Daam, M.A., Chelinho, S., Niemeyer, J., Owojori, O., de Silva, M., Sousa, J.P., van Gestel, C.A.M., Römbke, J. (2019). Environmental risk assessment of pesticides in tropical terrestrial ecosystems: current status and future perspectives. Ecotoxicology and Environmental Safety 181, 534-547.

Daam, M.A., Rico, A. (2016). Freshwater shrimps as sensitive test species for the risk assessment of pesticides in the tropics. Environmental Science and Pollution Research 25, 13235-13243.

Niemeyer, J.C., Moreira-Santos, M., Nogueira, M.A., Carvalho, G.M., Ribeiro, R., Da Silva, E.M., Sousa, J.P. (2010). Environmental risk assessment of a metal contaminated area in the Tropics. Tier I: screening phase. Journal of Soils and Sediments 10, 1557-1571.

Niva, C.C., Niemeyer, J.C., Rodrigues da Silva Júnior, F.M., Tenório Nunes, M.E., de Sousa, D.L., Silva Aragão, C.W., Sautter, K.D., Gaeta Espindola, E., Sousa, J.P., Römbke, J. (2016). Soil Ecotoxicology in Brazil is taking its course. Environmental Science and Pollution Research 23, 363-378.

Waichman, A.V., Römbke, J., Ribeiro, M.O.A., Nina, N.C.S. (2002). Use and fate of pesticides in the Amazon State, Brazil. Risk to human health and the environment. Environmental Science and Pollution Research 9, 423-428.

4.4.7. Question 1

What are the most important climatic factors affecting the fate and effects of chemicals when comparing temperate and tropical regions?

4.4.7. Question 2

Can tropical organisms be expected to be more sensitive to chemicals than temperate organisms? Please justify your answer.

4.4.7. Question 3

Should ecotoxicological test methods be adapted for their use in tropical regions? If yes, please provide two examples of adaptations that should be made.

4.4.7. Question 4

If a chemical is allowed for use in Europe, would you recommend its use in a tropical country without additional testing? Please justify your answer.

This page titled 2.3.4.4: Increasing ecological realism in toxicity testing is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.4: Increasing ecological realism in toxicity testing** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.4.5: Availability and bioavailability

3.6. Availability and bioavailability

3.6.1. Definitions

Authors: Martina Vijver

Reviewer: Kees van Gestel, Ravi Naidu

Leaning objectives:

You should be able to:

- understand that bioavailability consists of three principle processes.
- understand that bioavailability is a dynamic concept.
- understand why bioavailability is important to explain uptake and effects, essential for a proper risk assessment of chemicals.

Keywords: Chemical availability, actual and potential uptake, toxico-kinetics, toxico-dynamics.

Introduction:

Although many environmental chemists, toxicologists, and engineers claim to know what bioavailability means, the term eludes a consensus definition. Bioavailability may be defined as that fraction of chemical present in the environment that is or may become available for biological uptake by passage across cell membranes.

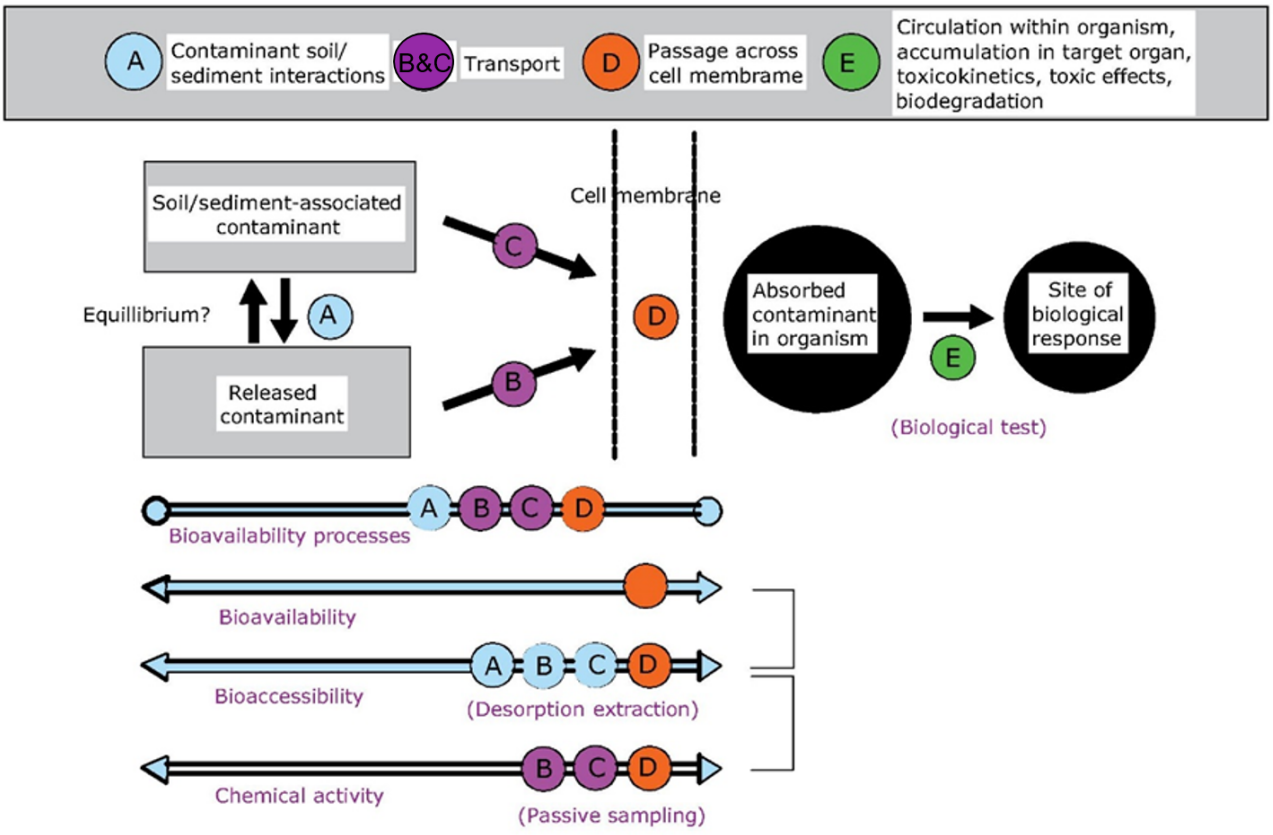


Figure 1. Bioavailability relates to a series of processes, ranging from processes external of organisms, towards internal tissues, and fully internal to the biological response site. Redrawn from Ortega-Calvo et al. (2015) by Wilma IJzerman.

Bioavailability generally is approached from a process-oriented point of view within a toxicological framework, which is applicable to all types of chemicals (Figure 1).

The first process is chemical availability which can be defined as the fraction of the total concentration of chemicals present in an environmental compartment that contributes to the exposure of an organism. The total concentration in an environmental compartment is not necessarily involved in the exposure, as a smaller or larger fraction of the chemical may be *bound to organic or inorganic components* of the environment. Organic matter and clay particles, for instance, are important in binding chemicals (see section on Soil), while also the presence of cations and pH are important factors modifying the partitioning of chemicals between different environmental phases (see section on Metal speciation).

The second process is the actual or potential uptake, described as the toxicokinetics of a substance which reflects the development with time of its concentration on, and in, the organism (see section on Bioconcentration and kinetics modelling).

The third process describes the internal distribution of the substance leading to its interaction(s) at the cellular site of toxicity activation. This is sometimes referred to as toxico-availability and also includes the biochemical and physiological processes resulting from the effects of the chemical at the site of action.

Details on the bioavailability concept described above as well as how the physico-chemical interactions influencing each process are described in the sections on Metal speciation and Bioconcentration and kinetics modelling.

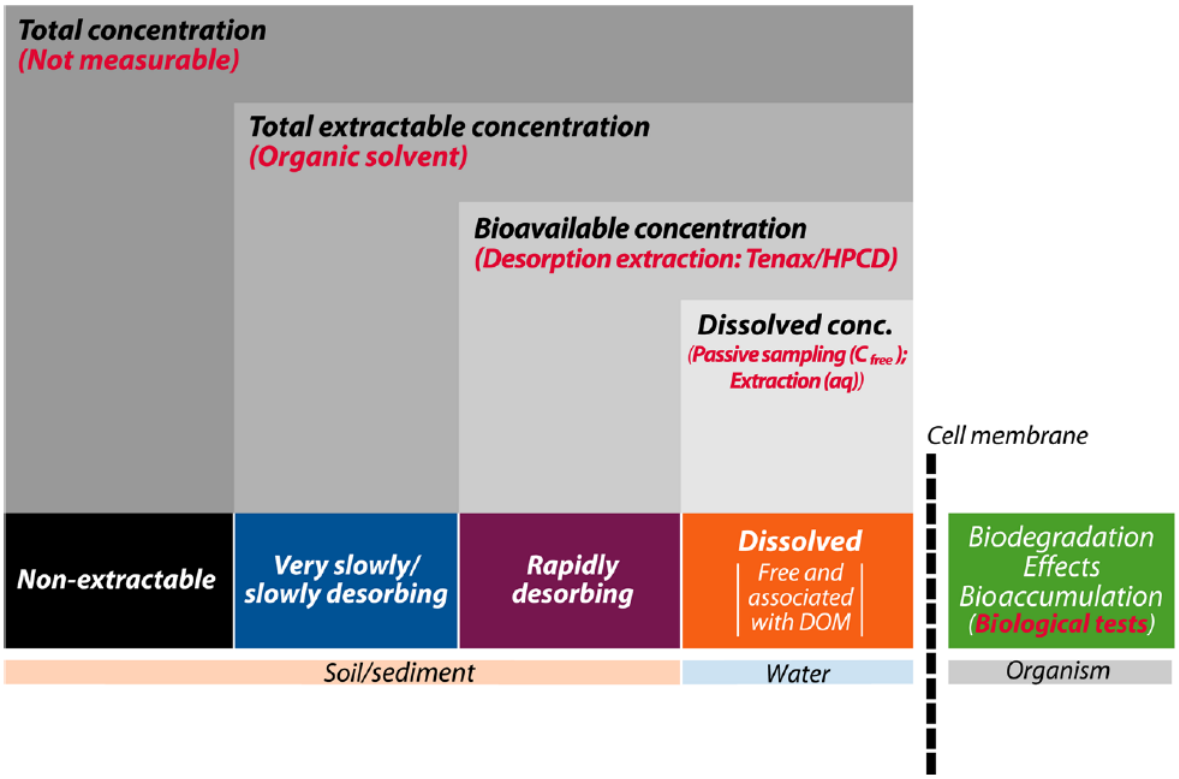


Figure 2. Bioavailability relates to a series of time frames, particularly in external processes (according to Ortega-Calvo et al. 2015).

Kinetics are involved in all three basic processes. The timeframe can vary from very brief (less than seconds) to very long in the order of hundreds of years. Figure 2 shows that some fractions of pollutants are present in soil or sediment, but may never

contribute to the transport of chemicals that could reach the internal site during an organism's lifespan. The fractions with different desorption kinetics may relate to different experimental techniques to determine the relevant bioavailability metric.

Box 1: Illustration of how bioavailability influences our human fitness

Iron deficiency occurs when a body has not enough iron to supply its needs. Iron is present in all cells of the human body and has several vital functions. It is a key component of the hemoglobin protein, carrying oxygen to the tissues from the lungs. Iron also plays an important role in oxidation/reduction reactions, which are crucial for the functioning of the cytochrome P450 enzymes that are responsible for the biotransformation of endogenic as well as xenobiotic chemicals. Iron deficiency therefore can interfere with these vital functions, leading to a lack of energy (feeling tired) and eventually to malfunctioning of muscles and the brain.

In case of iron deficiency, the medical doctor will prescribe Fe-supplements and iron-rich food such as red meat and green leafy vegetables like spinach. Although this will lead to a higher intake of iron (after all exposure is higher), it does not necessarily lead to a higher uptake as here bioavailability becomes important. It is advised to avoid drinking milk or caffeinated drinks together with eating iron-rich products or supplements because both drinks will prevent the absorption of iron in the intestinal tract. Calcium ions abundant in milk will compete with iron ions for the same uptake sites, so excess calcium will reduce iron uptake. Carbonates and caffeine molecules, but also phytate (inositol polyphosphate) present in vegetables, will strongly bind the iron, also reducing its availability for uptake.

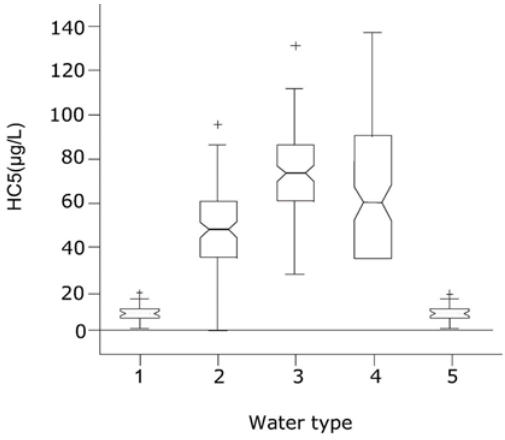


Figure 3. Bioavailability correction to estimate the HC5 copper concentration in relation to properties like dissolved organic carbon content and pH (Table 1) in order to estimate its risk in different water types (according to Vijver et al. 2008), in comparison to the 1.5 µg/L total dissolved Cu in surface waters as the current generic Dutch standard (horizontal line). Redrawn from Vijver et al. (2008) by Wilma Ijzerman

Bioavailability used in Risk Assessment

For regulatory purposes, it is necessary to use a straightforward approach to assess and prioritize contaminated sites based on their risk to human and environmental health. The bioavailability concept offers a scientific underpinned concept to be used in risk assessment. Examples for inorganic contaminants are the derived 2nd tier models such as the Biotic Ligand Models, while for organic chemicals the Equilibrium Partitioning (E_qP) concept (see Box 2 in the section on Sorption) is applied.

A quantitative example is given for copper in different water types in Figure 3 and Table 1, in which water chemistry is explicitly accounted for to enable estimating the available copper concentration. The current Dutch generic quality target for surface waters is 1.5 µg/L total dissolved copper. The bioavailability-corrected risk limits (HC5) for different water types, in most cases, exceeded this generic quality target.

Table 1. Bioavailability adjusted Copper 5% Hazardous Concentration (HC5, potentially affecting <5% of relevant species) for different water types.

Water type description	no.	DOC (mg/L)	pH	Average HC5 (µg/L)
Large rivers	1	3.1 ± 0.9	7.7 ± 0.2	9.6 ± 2.9
Canals, lakes	2	8.4 ± 4.4	8.1 ± 0.4	35.0 ± 17.9
Streams, brooks	3	18.2 ± 4.3	7.4 ± 0.1	73.6 ± 18.9
Ditches	4	27.5 ± 12.2	6.9 ± 0.8	64.1 ± 34.5
Sandy springs	5	2.2 ± 1.0	6.7 ± 0.1	7.2 ± 3.1

When the calculated HC5 value is lower, this means that the bioavailability of copper is higher and hence at the same total copper concentration in water the risk is higher. The bioavailability-corrected HC5s for Cu differ significantly among water types. The lowest HC5 values were found for sandy springs (water type V) and large rivers (water type I), which appear to be sensitive water bodies. These differences can be explained from partitioning processes (chemical availability) and competition processes (the toxicokinetics step) on which the BLMs are based. Streams and brooks (water type III) have rather high total copper concentrations without any adverse effects, which can be attributed to the protective effect of relatively high dissolved organic carbon (DOC) concentrations and the neutral to basic pH causing a high binding of Cu to the DOC.

For risk managers, this water type specific risk approach can help to identify the priority in cleanup activities among sites having elevated copper concentrations. It remains possible that, for extreme environmental situations (e.g., extreme droughts and low water discharges or extreme rain fall and high runoff), combinations of the water chemistry parameters may result in calculated HC5 values that are even lower than the calculated average values. For the latter (important) reason, the generic quality target is more strict.

References

Hamelink, J., Landrum, P.F., Bergman, H., Benson, W.H. (1994) Bioavailability: physical, chemical, and biological interactions, CRC Press.

Ortega-Calvo, J.J., Harmsen, J., Parsons, J.R., Semple, K.T., Aitken, M.D., Ajao, C., Eadsforth, C., Galay-Burgos, M., Naidu, R., Oliver, R., Peijnenburg, W.J.G.M., Römbke, J., Streck, G., Versonnen, B. (2015) From bioavailability science to regulation of organic chemicals. Environmental Science and Technology 49, 10255−10264.

Vijver, M.G., de Koning, A., Peijnenburg, W.J.G.M. (2008) Uncertainty of water type-specific hazardous copper concentrations derived with biotic ligand models. Environmental Toxicology and Chemistry 27, 2311-2319.

3.6.1. Question 1

What are the three processes that define bioavailability?

3.6.1. Question 2

Explain how metal uptake changes when high concentrations of dissolved organic matter are in the exposure medium.

3.6.1. Question 3

Explain why bioavailability is a dynamic concept.

3.6.2. Assessing available concentrations of organic chemicals

Author: Jose Julio Ortega-Calvo

Reviewers: John Parsons, Gerard Cornelissen

Learning objectives:

You should be able to:

- define the concept of freely dissolved concentrations and fast-desorbing fractions of organic chemicals in soil and sediment, as indicators of their bioavailability
- understand how to determine bioavailable concentrations with the use of passive sampling
- understand how to determine fast-desorbing fractions with desorption extraction methods.

Keywords: Bioavailability, Freely-dissolved concentration, Desorption, Passive sampling, Infinite sink

Introduction: Bioavailability through the water phase

In many exposure scenarios involving organic chemicals, ranging from a bacterial cell to a fish, or from a sediment bed to a soil profile, the organisms experience the pollution through the water phase. Even when this is not the case, for example when uptake is from sediment consumed as food, the aqueous concentration may be a good indicator of the bioavailable concentration, since ultimately a chemical equilibrium will be established between the solid phase, the aqueous phase (possibly in the intestine), and the organism. Thus, taking an aqueous sample from a given environment, and determining the concentration of a certain chemical with the appropriate analytical equipment seems a straightforward approach to assess bioavailability. However, especially for hydrophobic chemicals, which tend to remain sorbed to solid surfaces (see sections on Relevant chemical properties and Sorption of organic chemicals), the determination of the chemicals present in the aqueous phase, as a way to assess bioavailability, has represented a significant challenge to environmental organic chemistry. The phase exchange among different compartments often leads to equilibrium aqueous concentrations that are very low, because most of the chemicals remain associated to the solids, and after sustained exposure, to the biota. These freely dissolved concentrations (C_{free}) are very useful to determine bioavailability, as they represent the "tip of the iceberg" under equilibrium exposure, and are what organisms "see" (Figure 1, left). Similarly to the balance between gravity and buoyancy forces leading to iceberg flotation up to a certain level, C_{free} is determined by the equilibrium between sorption and desorption, and connected to the concentration of the sorbed chemical (C_{sorbed}) through a partitioning coefficient.

Biological uptake may also result in the fast removal of the chemical from the aqueous phase, and thus in further desorption from the solids, so equilibrium is never achieved, and actual aqueous concentrations are much lower than the equilibrium C_{free} (or even close to zero). In these situations, bioavailability is driven by the desorption kinetics of the chemical. Usually, desorption occurs as a biphasic process, where a fast desorption phase, occurring during a few hours or days, is followed by a much slower phase, taking months or even years. Therefore, for scenarios involving rapid exposures, or for studies on coupled desorption/biodegradation, the fast-desorbing fraction of the chemicals (F_{fast}) can be used to determine bioavailability. This fraction is often referred to as the bioaccessible fraction. Following the iceberg analogy (Figure 1, right), F_{fast} would constitute the upper iceberg fraction rapidly melting by sun irradiation, with a very minimal "visible" surface (representing the desorbed chemical in the aqueous solution, which is quickly removed by biological uptake). The slowly desorbing -or melting- fraction, F_{slow}, would remain in the sorbed state, within a given time span, having little interactions with the biota.

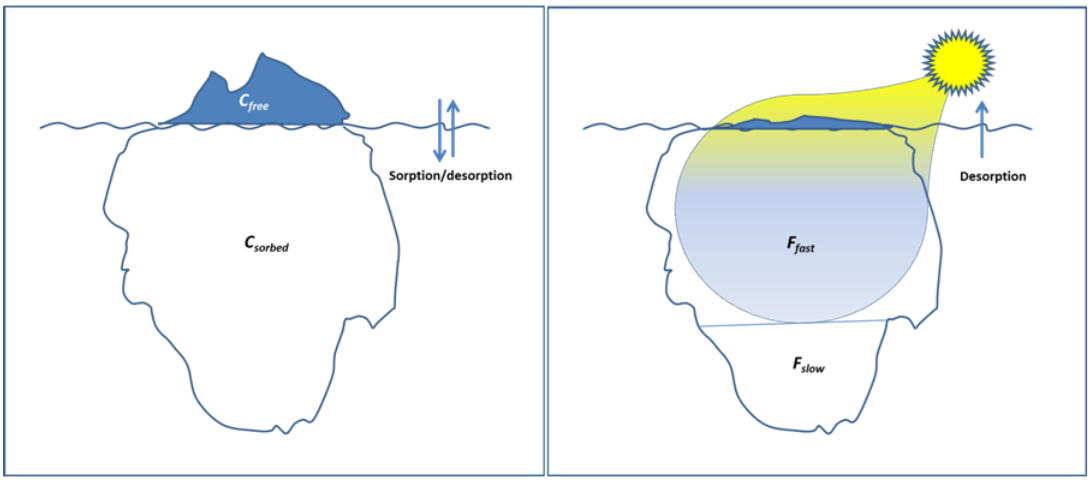


Figure 1. The magnitude of C_{free}, determined by the sorption/desorption equilibrium (similarly to a floating iceberg, left), can correspond to a minimal fraction of the total pollutant mass, but may constitute the main driver for bioavailability (and risk) in equilibrium exposure

scenarios. In non-equilibrium conditions (right, in analogy, a melting iceberg exposed to irradiation to sun), the fraction of sorbed chemical that can be rapidly mobilized, F_{fast}, can be taken as an estimate of bioavailability.

Determining bioavailability with passive sampling methods

C_{free} can be determined with a passive sampler, in the form of polymer-coated fibers or sheets (membranes) made of a variety of polymers, which establish an additional sorption equilibrium with the aqueous phase in contact with the soil or sediment (Jonker et al., 2018). Depending on the analytes of interest, different polymers, such as polydimethylsiloxane (PDMS) or polyethylene (PE), are used in passive samplers. The passive sampler, enriched in the analyte (similarly to the floating iceberg in Figure 1, left, where C_{sorbed} in this case is the concentration in the passive sampler), can be used in this way to determine indirectly the pollutant concentration present in the aqueous phase, even at very low concentrations, though the appropriate distribution ratio between sampler and water. In bioavailability estimations, passive sampling is designed for equilibrium and non-depletive conditions. This means that the amount of chemical sampled does not alter the solid-water equilibrium, i.e., it is essential that C_{free} is not affected significantly by the sampler. Equilibrium achievement is critical, and it may take days or weeks.

C_{free} can be calculated from the concentration of the pollutant in the passive sample polymer at equilibrium (C_p), and the polymer-to-water partitioning coefficient (K_{pw}):

$$C_{free} = \frac{C_p}{K_{pw}}$$

C_{free} values can be the basis of predictions for bioaccumulation that use the equilibrium partitioning approach, either directly or through a bioconcentration factor, and for sediment toxicity in conjunction with actual toxicity tests. Passive sampling methods are well suited for contaminated sediments, and they have already been implemented in regulatory environmental assessments based on bioavailability (Burkhard et al., 2017).

Determining bioavailability with desorption extraction methods

The determination of F_{fast} can be achieved with the use of methods that trap the desorbed chemical once it appears in the aqueous phase. Far from equilibrium conditions, desorption is driven to its maximum rate by placing a material in the aqueous phase that acts as an infinite sink (comparable to the sun irradiation of a melting iceberg in Figure 1, right). The most accepted materials for these desorption extraction methods are Tenax, a sorptive resin, and cyclodextrin, a solubilizing agent (ISO, 2018). These methods allow a permanent aqueous chemical concentration of almost zero, and therefore, sorption of the chemical back to the soil or sediment can be neglected. Several extraction steps can be used, covering a variable time span, which depends on the environmental sample.

The following first-order, two-compartment kinetic model can be used to analyze desorption extraction data:

$$\frac{St}{S0} = F_{fast}\, e^{-k_{fast}\, t} + F_{slow}\, e^{-k_{slow}\, t}$$

In this equation, S_t and S_o (mg) are the soil-sorbed amounts of the chemical at time t (h) and at the start of the experiment, respectively. F_{fast} and F_{slow} are the fast- and slow-desorbing fractions, and $k_{fast}$and k_{slow}(h^{-1}) are the rate constants of fast and slow desorption, respectively. To calculate the values of the different constants and fractions (F_{fast}, F_{slow}, k_{fast}, and k_{slow}) exponential curve fitting can be used. The ln form of the equation can be used to simplify curve fitting.

Once the desorption kinetics are known, the method can be simplified for a series of samples, by using single time point-extractions. A time period of 20 h has been suggested as a sufficient time period to approximate F_{fast}. It is highly convenient for operational reasons (ISO, 2018), but indicative at best, since the time needed to extract Ffast tends to vary between chemicals and soils/sediments.

References

Burkhard, L.P., Mount, D.,R., Burgess, R.,M. (2017). *Developing Sediment Remediation Goals at Superfund Sites Based on Pore Water for the Protection of Benthic Organisms from Direct Toxicity to Nonionic Organic Contaminants* EPA/600/R 15/289; U.S. Environmental Protection Agency Office of Research and Development: Washington, DC.

ISO (2018). Technical Committee ISO/TC 190 *Soil quality - Environmental availability of non-polar organic compounds - Determination of the potentially bioavailable fraction and the non-bioavailable fraction using a strong adsorbent or complexing*

agent; International Organization for Standardization: Geneva, Switzerland.

Jonker, M.T.O., van der Heijden, S.A., Adelman, D., Apell, J.N., Burgess, R.M., Choi, Y., Fernandez, L.A., Flavetta, G.M., Ghosh, U., Gschwend, P.M., Hale, S.E., Jalalizadeh, M., Khairy, M., Lampi, M.A., Lao, W., Lohmann, R., Lydy, M.J., Maruya, K.A., Nutile, S.,A., Oen, A.M.P., Rakowska, M.I., Reible, D., Rusina, T.P., Smedes, F., Wu, Y. (2018) Advancing the use of passive sampling in risk assessment and management of sediments contaminated with hydrophobic organic chemicals: results of an international ex situ passive sampling interlaboratory comparison. Environmental Science & Technology *52* (6), 3574-3582.

3.6.2. Question 1

Describe why the freely dissolved concentration is often used as an indicator of the bioavailability of organic chemicals in soil and sediment.

3.6.2. Question 2

What is passive sampling and how can this be used to determine the bioavailable concentrations of chemicals in soil or sediment?

3.6.2. Question 3

What is the meant by fast-desorbing fraction of a chemical in soil or sediment and how can this be determined with desorption extraction methods?

3.6.3. Assessing available metal concentrations

Authors: Kees van Gestel

Reviewer: Martina Vijver, Steve Lofts

Leaning objectives:

You should be able to:

- mention different methods for assessing chemically available metal fractions in soils and sediments.
- indicate the relative binding strengths of metals extracted with the different methods or in different steps of a sequential extraction procedure.
- explain the pros and cons of chemical extraction methods for assessing metal (bio)availability in soils and sediments.

Keywords: Chemical availability, actual and potential uptake, toxicokinetics, toxicodynamics.

Introduction:

Total concentrations are not very informative about the availability of metals in soils or sediments. Fate and behavior of metals - in general terms mobility - as well as biological uptake and toxicity is highly determined by their speciation. Speciation describes the partitioning of a metal among the various forms in which it may exist (see section on Metal speciation). For assessing the risk of metals to man and the environment, speciation therefore is highly relevant as it may determine their availability for uptake and effects in organisms. Several tools have been developed to determine available metal concentrations or their speciation in soils and sediments. As indicated in the section on Availability and bioavailability, such chemical methods are just indicative, and to a large extent ignore dynamics of availability. Moreover, availability is also influenced by biological processes, with abiotic-biotic interactions influencing the bioavailability process being species- and often even life-stage specific. Nevertheless, chemical extractions may provide useful information to predict or estimate the potential risks of metals and therefore are preferred over the determination of total metal concentrations.

The available methods include:

1. Porewater extraction
2. Extractions with water
3. Extractions with diluted salts
4. Extractions with chelating agents
5. Extractions with diluted acids
6. Sequential extractions using a series of different extraction solutions

7. Passive sampling methods

Porewater extraction probably best approaches the readily available fraction of metals in soil, which drives mobility and is the fraction of metals experienced directly by many organisms exposed. In general, pore water is extracted from soil or sediment by centrifugation, and filtration over a 0.45 µm (or 0.22 µm) filter to remove larger particles and perhaps some of the dissolved organic matter. Filtration, however, will not remove all complexes, making it impossible to determine solely the dissolved metal fraction in the pore water. Nevertheless, porewater metal concentrations have been shown to have significant correlations with metal uptake (e.g. for copper uptake by barley and tomato by Zhao et al., 2006) and to be useful for predicting toxic threshold concentrations of metals, with correction for pH (e.g. for nickel toxicity to tomato and barley by Rooney et al., 2007).

Extraction with water simulates the immediately available fraction, so the fraction present in the soil solution or pore water. By extracting soil with water, the pore water however, is diluted, which on one hand may facilitate metal analysis by creating larger volumes of solution, but on the other hand may lead to differences between measured and actual metal concentrations in the pore water as it may impact chemical equilibria.

Extraction with diluted salts aims to determine the fraction of metal that is easily available or may become available as it is in the exchangeable form. This refers to cationic metals that may be bound to the negatively charged soil particles (see section on Soil). Buffered salt solutions, for instance 1 M NH_4-acetate at pH 4.8 (with acetic acid) or at pH 7, may under- or overestimate available metal concentrations because of their interference with soil pH. Unbuffered salt solutions therefore are more widely used and may for instance include 0.001 or 0.01 M $CaCl_2$, 0.1 M $NaNO_3$ or 1 M NH_4NO_3 (Gupta and Aten, 1993; Novozamsky et al., 1993). Gupta and Aten (1993) showed good correlations between the uptake of some metals in plants and 0.1 M $NaNO_3$ extractable concentrations in soil, while Novozamsky et al. (1993) found similar well-fitting correlations using 0.01 M $CaCl_2$. The latter method also seemed well capable of predicting metal uptake in soil invertebrates, and therefore has been more widely accepted for predicting metal availability in soil ecotoxicology. Figure 1 (Zhang et al., 2019) provides an example with the correlation between Pb toxicity to enchytraeid worms in different soils and 0.01 M $CaCl_2$ extractable concentrations.

Extractions with water (including porewater) and dilute salts are most accurately described as measures of the chemical solubility of the metal in the soil. The values obtained can be useful indicators of the relative metal reactivity across soils, but tend to be less useful for bioavailability assessment, unless the soils under consideration have a narrow range of soil properties. This is because the solutions obtained from such soils themselves have varying chemical properties (e.g. pH, DOC concentration) which are likely to affect the availability of the measured metal to organisms.

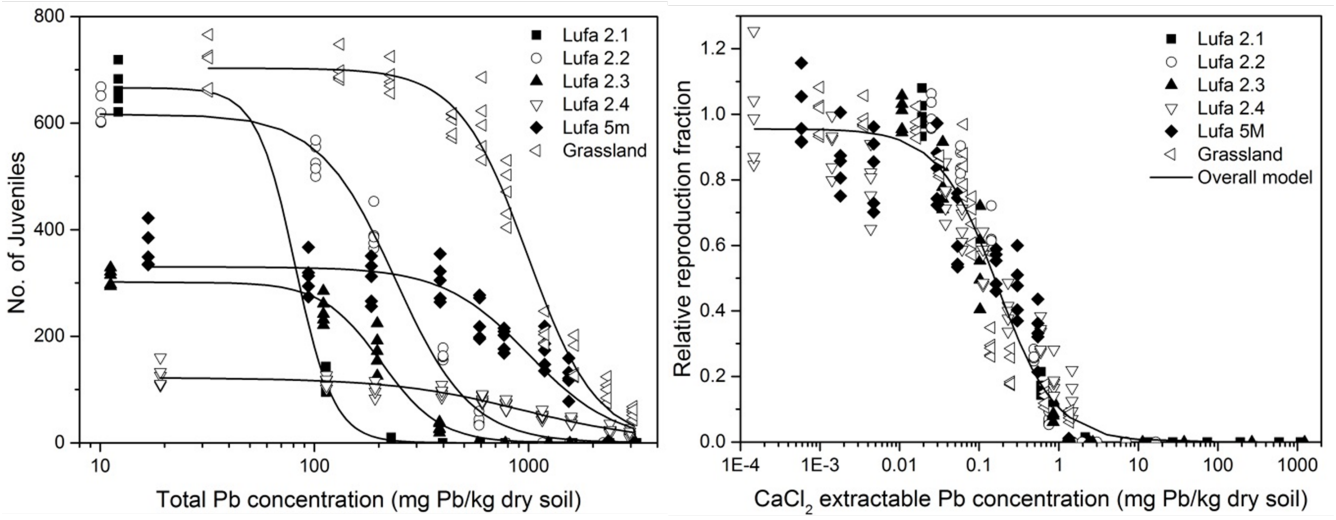


Figure 1. Effects of $Pb(NO_3)_2$ on the reproduction of Enchytraeus crypticus after three weeks exposure in six natural soils. Pb concentrations are expressed as total (A) and 0.01 M $CaCl_2$ extractable concentrations in soil (B). Lines show the fit of a logistic dose-response curve. When expressed on the basis of 0.01 M $CaCl_2$ extrable concentrations, dose-response curves did not significantly differ and a single curve is shown. Data taken from Zhang et al. (2019).

Extraction with chelating agents, such as EDTA (0.01-0.05 M) or DTPA (0.005 M) (as their sodium or ammonium salts), aims at assessing the availability of metals for plants. Many plants have the ability to actively affect metal speciation in the soil by producing root exudates. These extractants may form very stable water-soluble complexes with many different polyvalent cationic metals. It should be noted that the large variation in plant species and corresponding physiologies as well as their interactions with symbiotic microorganisms (e.g. mycorrhizal fungi) make that there is no single extraction method is capable of properly predicting metal availability to all plant species.

Extraction with diluted acids has been advocated for predicting the potentially available fraction of metals in soils, so the fraction that may become available in the long run. It is a quite rigorous extraction method that can be executed in a robust way. Metal concentrations determined by extracting soils with 0.43 M HNO_3 showed very good correlation with oral bioaccessible concentrations (Rodrigues et al., 2013), probably because it to some degree simulates metal release under acidic stomach conditions.

Both extraction methods with chelating agents and diluted acid may also dissolve solids, such as carbonates and Fe- and Al-oxides. This raises concerns as to the interpretation of results of these extraction systems, and especially to their generalization to different soil-plant systems (Novozamsky et al., 1993). The extractions with chelating agents and dilute acids are considered methods to estimate the 'geochemically active' metal in soil - the pool of adsorbed metal that can participate in solid-solution adsorption/desorption and exchange equilibria on timescales of hours to days. This pool, along with the basic soil properties such as pH etc., also controls the readily available concentrations obtained with water/weak salt/porewater extraction. From the bioavailability point of view, these extractions tend to be most useful as inputs to bioavailability/toxicity models such as that of Lofts et al. (2014), which take further account of the effects of metal speciation and soil chemistry on metal bioavailability to environmental organisms.

Sequential extraction brings together different extraction methods, and aims to determining either how strongly metals are retained or to which components of the solid phase they are bound in soils or sediments. This allows to determine how metals are bound to different fractions within the same soil or sediment, and allows interpretation to the bioavailability dynamics. By far the most widely used method of sequential extraction is the one proposed by Tessier et al. (1979). Five fractions are distinguished, indicating how metals are interacting with soil or sediment components: see Figure 2.

Where the Tessier method aims at assessing the environmental availability of metals in soils and sediments, similar sequential extraction methods have also been developed for assessing the potential availability of metals for humans (bioaccessibility) following gut passage of soil particles (see e.g. Basta and Gradwohl, 2000).

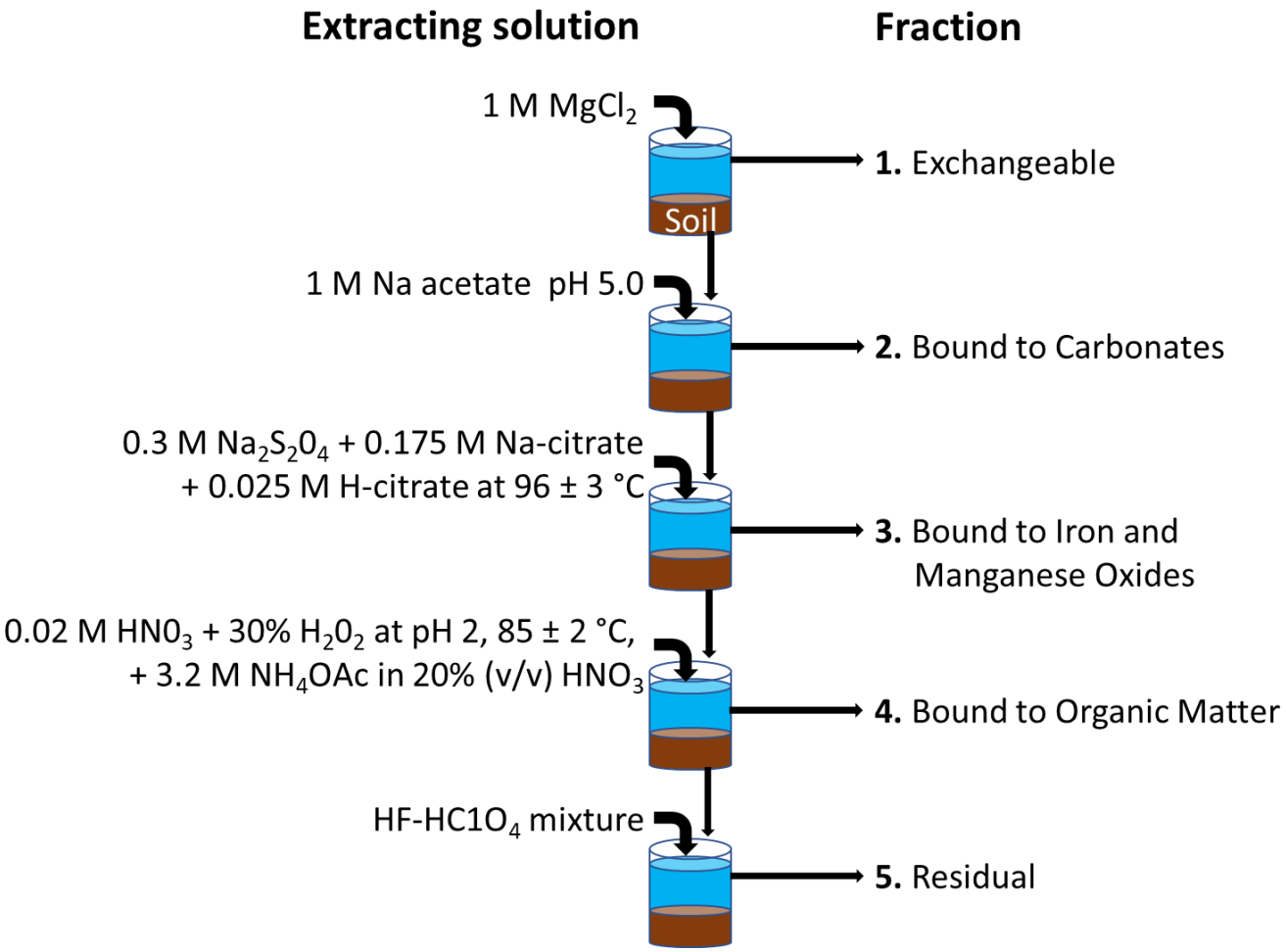


Figure 2. Schematic presentation of the sequential extraction of soil or sediment samples following the method of Tessier et al. (1979). The fractions obtained give an indication of the sites where metals are bound in the soil or sediment, and represent also an increasing binding strength, going from exchangeable to residual. Source: Kees van Gestel.

Passive sampling may also be applied to assess available metal concentrations. The best known method is that of Diffusive Gradients in Thin films (DGT), developed by Zhang et al. (1998). In this method, a resin (Chelex) with high affinity for metals is placed in a device and covered with a diffusive gel and a 0.45 µm cellulose nitrate membrane (Figure 3). The membrane is brought into contact with the soil. Metals dissolved in the soil solution will diffuse through a membrane and diffusive gel and bind to the resin. Based on the thickness of the membrane and gel and the contact time with the soil, the metal concentration in the pore water can be calculated from the amount of metal accumulated in the resin. The method may be indicative of available metal concentrations in soils and sediments, but can only work effectively when soil is sufficiently moist to guarantee optimal diffusion of metals to the resin. For the same reasons, the method probably is better suited for assessing the availability of metals to plants than to invertebrates, especially for animals that are not in continuous contact with the soil solution.

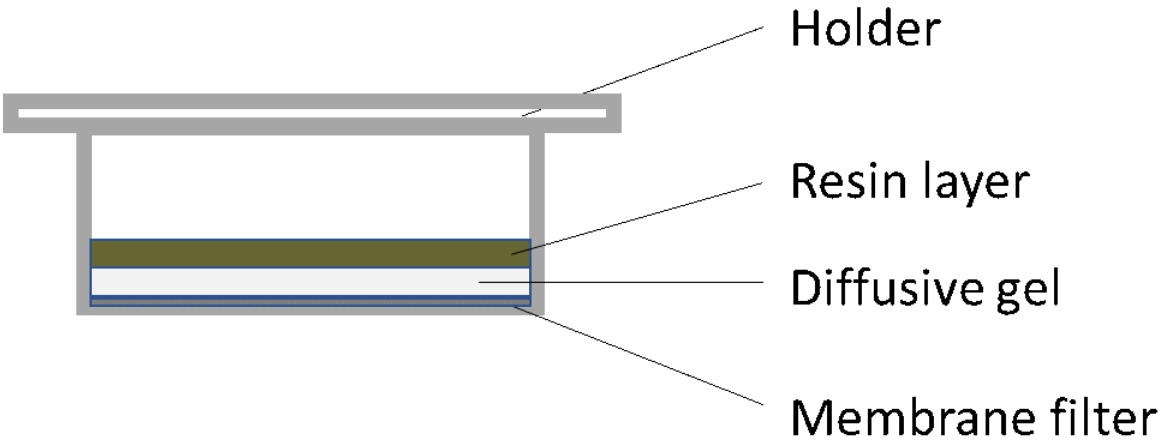


Figure 3.
Device used in the Diffusive Gradients in Thin film (DGT) method for determining available metal concentrations in soil and sediment (adapted from Zhang et al., 1998). The device is placed on the soil or sediment in such a way that the membrane filter makes contact with the porewater. Metals may diffuse from the porewater to the resin layer. See text for further explanation.

Several of the above described methods have been adopted by the International Standardization Organization (ISO) in (draft) standardized test guidelines for assessing available metal fractions in soils, sediments and waste materials, e.g. to assess the potential for leaching to groundwater or their potential bioaccessibility. This includes e.g. ISO/TS 21268-1 (2007) "Soil quality - Leaching procedures for subsequent chemical and ecotoxicological testing of soil and soil materials - Part 1: Batch test using a liquid to solid ratio of 2 l/kg dry matter", ISO 19730 (2008) "Soil quality -Extraction of trace elements from soil using ammonium nitrate solution" and ISO 17586 (2016) "Soil quality -- Extraction of trace elements using dilute nitric acid".

References:

Basta, N., Gradwohl, R. (2000). Estimation of Cd, Pb, and Zn bioavailability in smelter-contaminated soils by a sequential extraction procedure. Journal of Soil Contamination 9, 149-164.

Gupta, S.K., Aten, C. (1993). Comparison and evaluation of extraction media and their suitability in a simple model to predict the biological relevance of heavy metal concentrations in contaminated soils. International Journal of Environmental Analytical Chemistry 51, 25-46.

Lofts, S., Spurgeon, D.J., Svendsen, C., Tipping, E. (2004). Deriving soil critical limits for Cu, Zn, Cd, and Pb: A method based on free ion concentrations. Environmental Science and Technology 38, 3623-3631.

Novozamsky, I., Lexmond, Th.M., Houba, V.J.G. (1993). A single extraction procedure of soil for evaluation of uptake of some heavy metals by plants. International Journal of Environmental Analytical Chemistry 51, 47-58.

Rodrigues, S.M., Cruz, N., Coelho, C., Henriques, B., Carvalho, L., Duarte, A.C., Pereira, E., Römkens, P.F. (2013). Risk assessment for Cd, Cu, Pb and Zn in urban soils: chemical availability as the central concept. Environmental Pollution 183, 234-242.

Rooney, C.P., Zhao, F.-J., McGrath, S.P. (2007). Phytotoxicity of nickel in a range of European soils: Influence of soil properties, Ni solubility and speciation. Environmental Pollution 145, 596-605.

Tessier, A., Campbell, P.G.C., Bisson, M. (1979). Sequential extraction procedure for the speciation of particulate trace metals. Analytical Chemistry 51, 844-851.

Zhang, H., Davison, W., Knight, B., McGrath, S. (1998). In situ measurements of solution concentrations and fluxes of trace metals in soils using DGT. Environmental Science and Technology 32, 704-710.

Zhang, L., Verweij, R.A., Van Gestel, C.A.M. (2019). Effect of soil properties on Pb bioavailability and toxicity to the soil invertebrate Enchytraeus crypticus. Chemosphere 217, 9-17.

Zhao, F.J., Rooney, C.P., Zhang, H., McGrath, S.P. (2006). Comparison of soil solution speciation and diffusive gradients in thin-films measurement as an indicator of copper bioavailability to plants. Environmental Toxicology and Chemistry 25, 733-742.

3.6.3. Question 1

Which metal fractions are extracted with the Tessier method, and what does that say about metal availability?

3.6.3. Question 2

Why is porewater extraction to be preferred above water extraction?

3.6.3. Question 3

What is the principle of the DGT passive sampling method, and why does it not work for fairly dry soils?

3.6.3. Question 4

Which method may give the best estimate of potentially available metals for oral uptake by mammals?

3.6.3. Question 5

What are the pros and cons of chemical extraction methods for assessing the availability of metals in soils or sediments

---

This page titled 2.3.4.5: Availability and bioavailability is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.5: Availability and bioavailability** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.4.6: Degradation

3.7. Degradation

3.7.1. Chemical and photochemical degradation processes

Authors: John Parsons

Reviewers: Steven Droge, Kristopher McNeill

Leaning objectives:

You should be able to:

- understand the role of chemical and photochemical reactions in the removal of organic chemicals from the environment
- understand the most important chemical and photochemical reactions in the environment
- understand the role of direct and indirect photodegradation

Keywords: Environmental degradation reactions, Hydrolysis, Reduction, Dehalogenation, Oxidation, Photodegradation

Introduction

Transformation of organic chemicals in the environment can occur by a variety of reactions. These may be purely chemical reactions, such as hydrolyses or redox reactions, photochemical reactions with the direct or indirect involvement of light, or biochemical reactions. Such transformations can change the biological activity (toxicity) of a molecule; it can change the physico-chemical properties and thus change its environmental partitioning processes; it can change its bioavailability, for example facilitating biodegradation; or it may contribute to the complete removal (mineralization) of the chemical from the environment. In many cases, chemicals may be removed by combinations of these different processes and it is sometimes difficult to unequivocally identify the contributions of the different mechanisms. Indeed, combinations of different mechanisms are sometimes important, for example in cases where microbial activity is responsible for creating conditions that favour chemical reactions. Here we will focus on two types of reactions: Abiotic (dark) reactions and photochemical reactions. Biodegradation reactions are covered elsewhere (see section on Biodegradation).

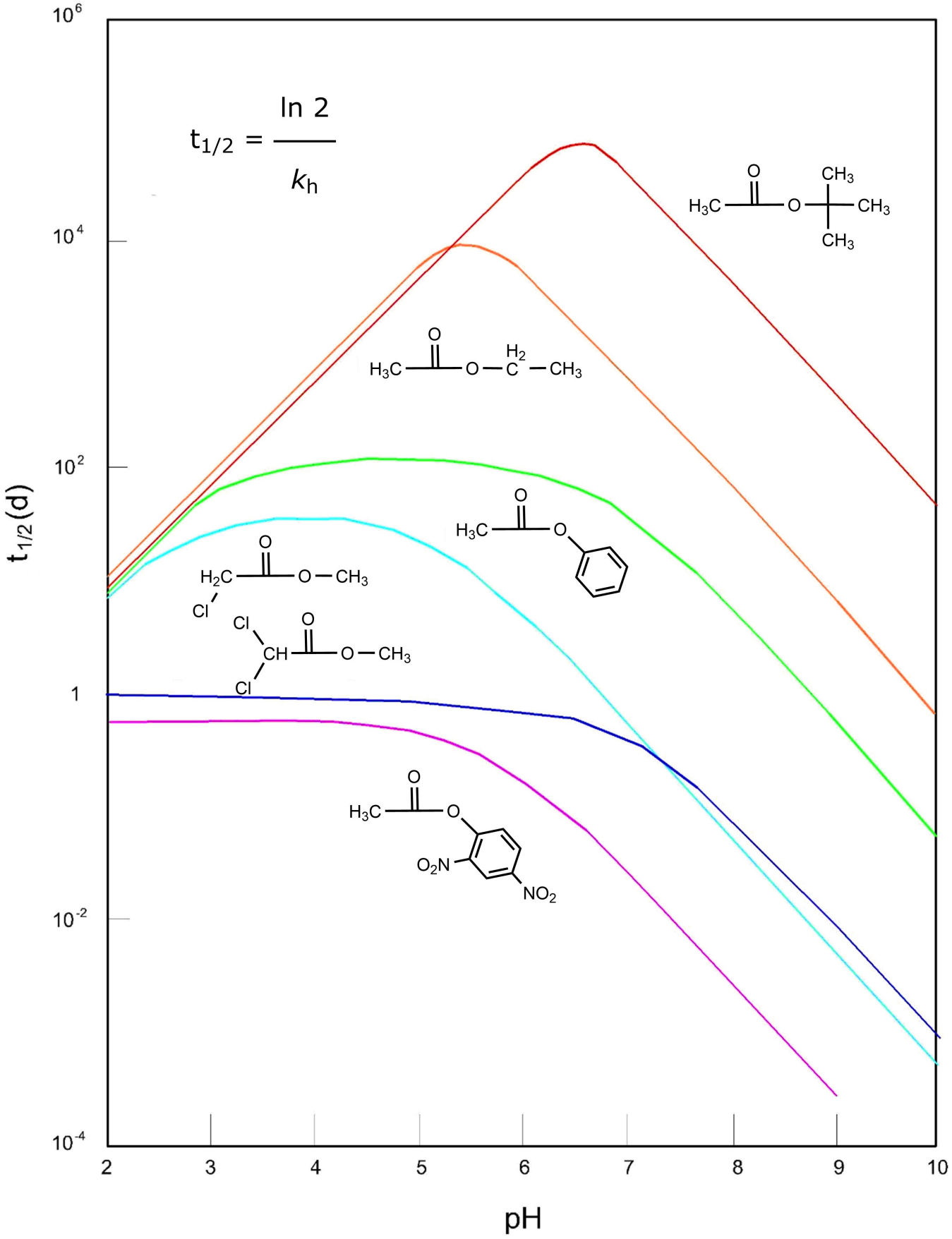
Chemical degradation

Hydrolytic reactions are important chemical reactions removing organic contaminants and are particularly important for chemicals containing acid derivatives as functional groups. Common examples of such chemicals are pesticides of the organophosphate and carbamate classes such as parathion, diazinon, aldicarb and carbaryl. Organophosphate chemicals are also used as flame retardants and are widely distributed in the environment. Some examples of hydrolysis reactions are shown in Figure 1.

Figure 1 Examples of hydrolyses of esters and carbamates (redrawn after Van Leeuwen and Vermeire, 2007).

As the name suggests, hydrolysis reactions involve using water (*hydro-*) to break (*-lysis*) a bond. Hydrolyses are reactions with water to produce an acid and either an alcohol or amine as products. Hydrolyses can be catalysed by either OH^- or H^+ ions and their rates are therefore pH dependent. Some examples of pH-dependent ester hydrolysis reactions are shown in Figure 2.

Halogenated organic molecules may also be hydrolysed to form alcohols (releasing the halogen as a halide ion). The rates of these reactions vary strongly depending on the structure of the organohalogen molecule and the halogen substituent (with Br and I being substituted more rapidly than Cl, and much more rapidly than F) and in general the rates of these reactions are too slow to be of more than minor importance except for tertiary organohalogens and secondary organohalogens with Br and I (Schwarzenbach et al. 2017).



https://bio.libretexts.org/@go/page/108683

Figure 2. Examples of pH dependent ester hydrolysis reactions (Schwarzenbach et al. 2017). Note that the y-axis is half-life (on a logarithmic scale), meaning high values correspond to slow reactions. Redrawn by Wilma Ijzerman.

In some cases, other substitution reactions not involving water as reactant may be important. Some examples include Cl^- in seawater converting CH_3I to CH_3Cl and reaction of thiols with alkyl bromines in anaerobic groundwater and sediment porewater under sulfate-reducing conditions (Schwarzenbach et al. 2017)

Redox (reduction and oxidation) reactions are another important reaction class involved in the degradation of organic chemicals. In the presence of oxygen, the oxidation of organic chemicals is thermodynamically favourable but occurs at insignificant rates unless oxygen is activated in the form of oxygen radicals or peroxides (following light absorption for example, see below) or if the reaction is catalysed by transition metals or transition metal-containing enzymes (see the sections on Biodegradation and Xenobiotic metabolism and defence).

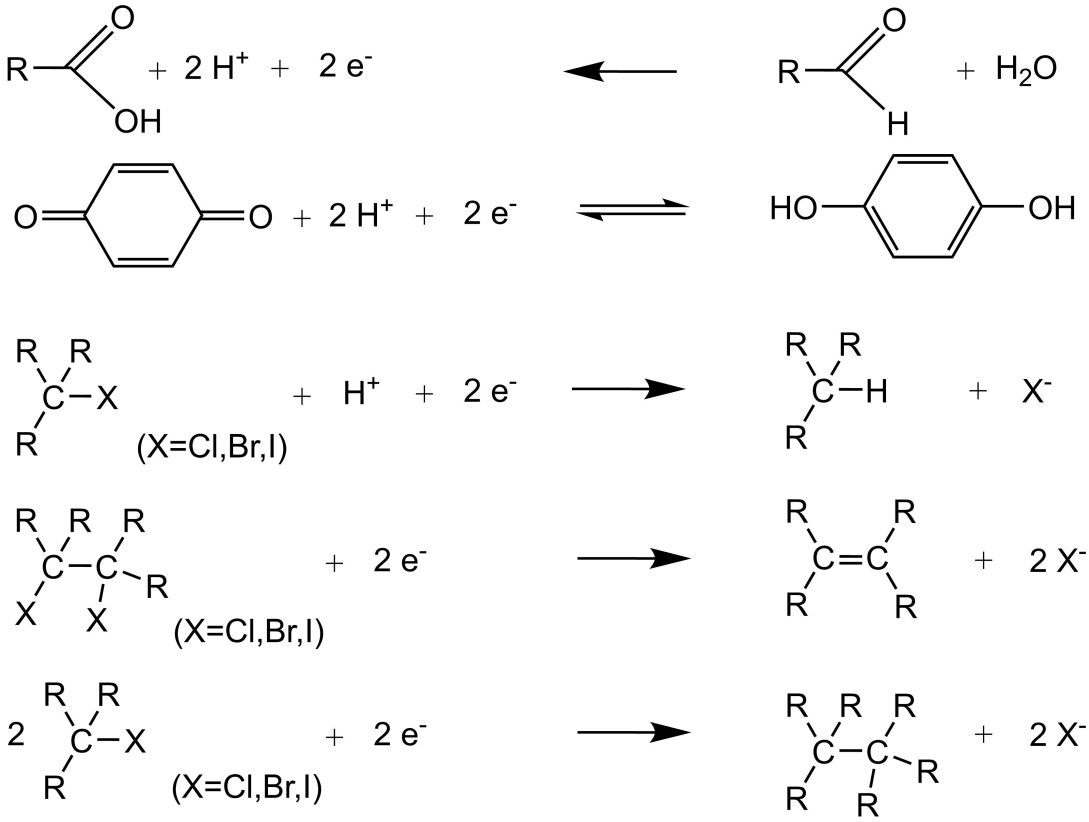
Reduction reactions are important redox reactions for environmental contaminants in anaerobic environments such as sediment and groundwater aquifers. Under these conditions, organic chemicals containing reducible functional groups such as carboxylic acids and nitro groups undergo reduction reactions (Table 1).

Table 1: Examples of chemical redox reactions that may occur in the environment (adapted from Schwarzenbach et al. 2017)

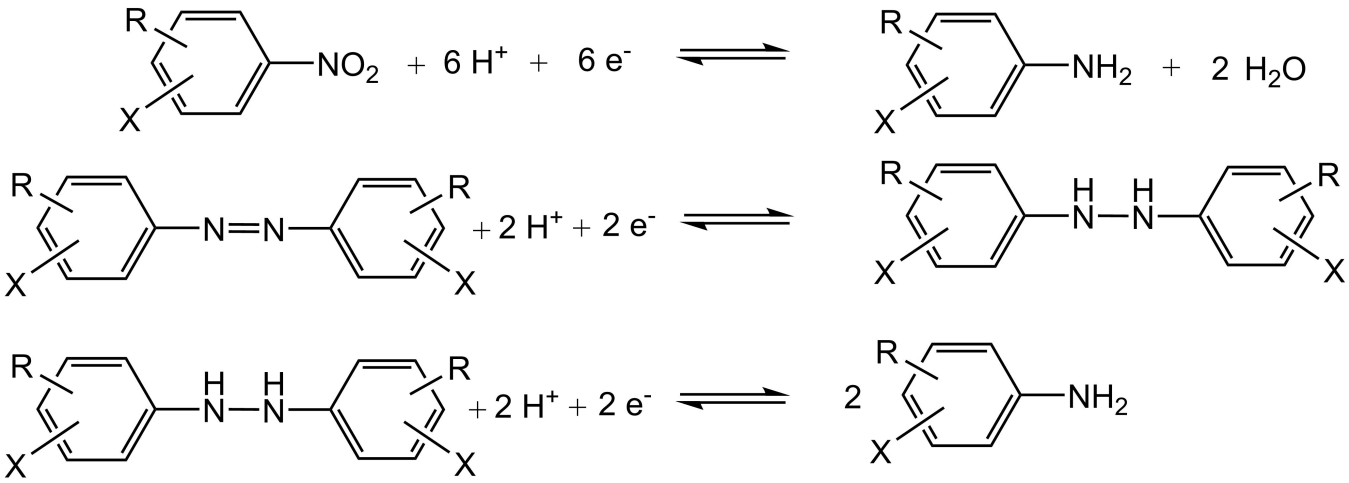
Oxidized species ⇌ **Reduced species**

Reduction (⇌) **Oxidation**

Change in Oxidation state of Carbon Atom(s)



Change in Oxidation state of Nitrogen Atom(s)



Change in Oxidation state of Nitrogen Atom(s)



Figure 3 shows reductive dehalogenation reactions of hexachlorobenzene.

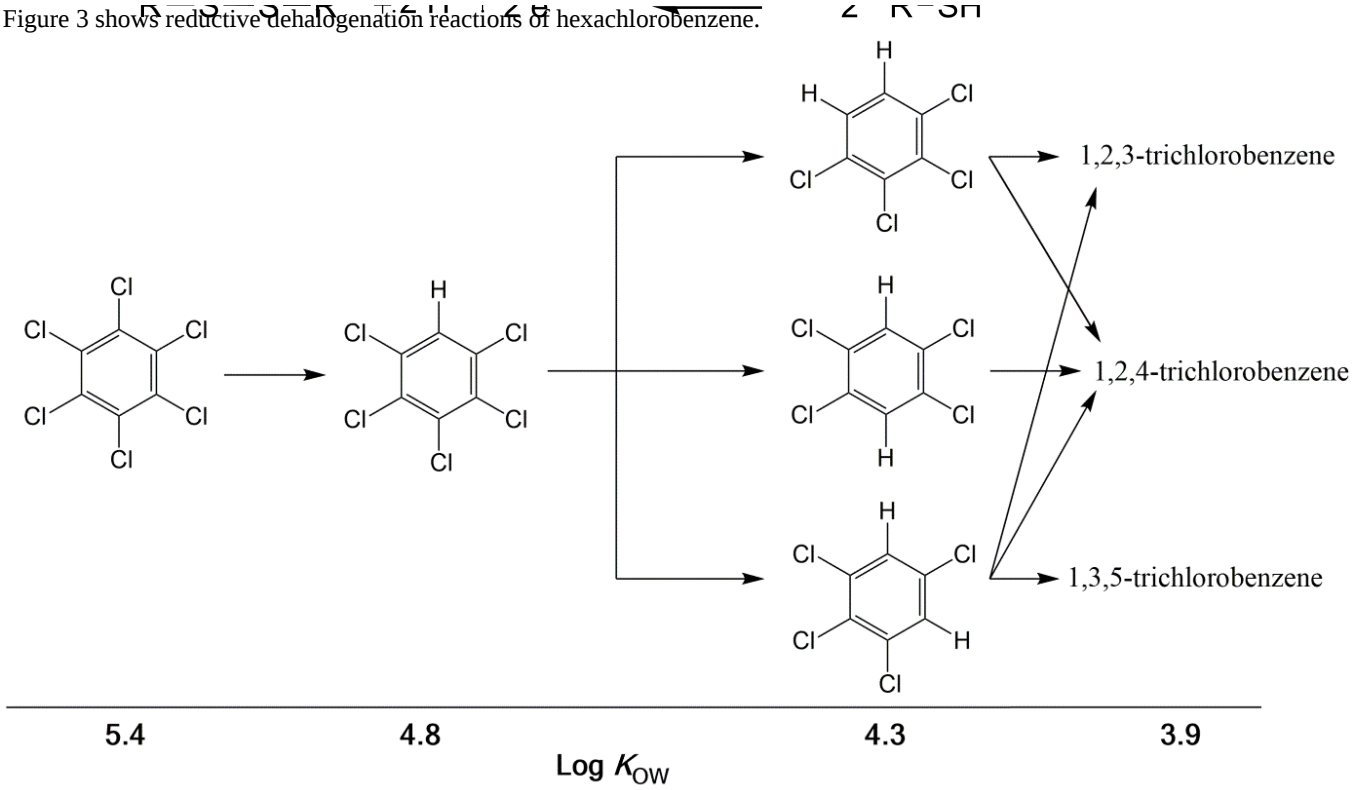


Figure 3. Reductive dehalogenation of hexachlorobenzene to less hydrophobic dechlorinated products (redrawn after Van Leeuwen and Vermeire, 2007).

Photodegradation

Sunlight is an important source of energy to initiate chemical reactions and photochemical reactions are particularly important in the atmosphere. Aromatic compounds and other chemicals containing unsaturated bonds that are able to absorb light in the frequency range available in sunlight become exited (energized) and this can lead to chemical reactions. These reactions lead to cleavage of bonds between carbon atoms and other atoms such as halogens to produce radical species. These radicals are highly reactive and react further to remove hydrogen or OH radicals from water to produce C-H or C-OH bonds or may react with themselves to produce larger molecules. Well known examples of atmospheric photochemical stratospheric reactions of CFCs that have had a negative impact on the so-called ozone layer and photochemical oxidations of hydrocarbons that are involved in the generation of smog.

In the aquatic environment, light penetration is sufficient to lead to photochemical reactions of organic chemicals at the water surface or in the top layer of clear water. The presence of particles in a waterbody reduces light intensity through light scattering as does dissolved organic matter through light absorption. Photodegradation contributes significantly to removing oil spills and appears to favour the degradation of longer chain alkanes compared to the preferential attack of linear and small alkanes by biodegradation (Garrett et al., 1998). Cycloalkanes and aromatic hydrocarbons are also removed by photodegradation (D'Auria et al., 2009). There is comparatively little known about the role photodegradation of other organic pollutants in the marine environment although there is, for example, evidence that triclosan is removed by photolysis in the German Bight area of the North Sea (Xie et al., 2008). In the soil environment, there is some evidence that photodegradation may contribute to the removal of a variety of organic chemicals such as pesticides and chemicals present in sewage sludge that is used as a soil amendment but the significance of this process is unclear. Similarly, chemicals that have accumulated in ice, for example as a result of long range transport to polar regions, also seem to be susceptible to photodegradation. Some examples of photodegradation reactions are shown in Figure 4.

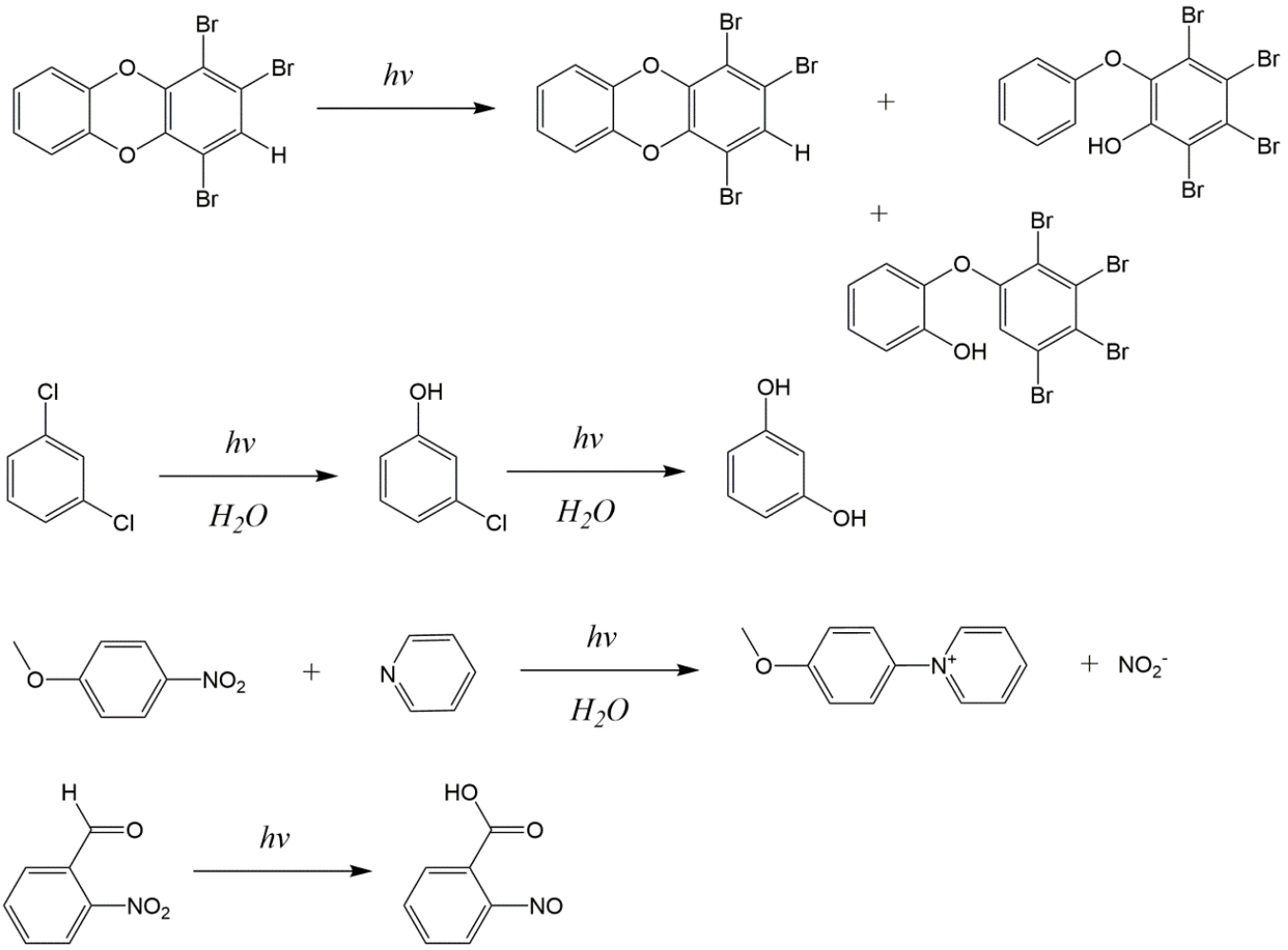


Figure 4. Some examples of photodegradation reactions (redrawn after Van Leeuwen and Vermeire, 2007). (Steven Droge 2019)

An important category of photochemical reactions are indirect reactions in which organic chemicals react with photochemically produced radicals, in particular with reactive oxygen species such as OH radicals, ozone and singlet oxygen. These reactive species are present at very low concentrations but are so reactive that under certain conditions they can contribute significantly to the removal of organic chemicals. Products of these reactions are a variety of oxidized derivatives which are themselves radicals and therefore react further. OH radicals are the most important of these photochemically produced species and can react with organic chemicals by removing hydrogen radicals, reacting with unsaturated bonds in alkenes, aromatics etc. to produce hydroxylated products. In water, natural organic matter absorbs light and can participate in indirect photodegradation reactions. Other constituents in surface water, such as nitrogen oxides and iron complexes may also be involved in indirect photodegradation reactions.

References

Schwarzenbach, R.P., Gschwend, P.M., Imboden, D.M. (2017). Environmental Organic Chemistry, Third Edition, Wiley, ISBN 978-1-118-76723-8

van Leeuwen, C.J., Vermeire, T.G. (2007). Risk Assessment of Chemicals: An Introduction (2nd ed.), Springer, ISBN 978-1-4020-6101-1

3.7.1. Question 1

Which organic chemicals would you expect to undergo hydrolytic degradation in the environment? Explain why these reaction depend on the pH.

3.7.1. Question 2

Reductive dehalogenation reactions are often observed to occur for organochlorine compounds in anaerobic environments. Why are these reactions called reductive dehalogenation? What products would you expect to be formed by the reductive dehalogenation of tetrachloroethene ($Cl_2C=CCl_2$)? Describe two ways in which bacteria are involved in these reactions.

3.7.1. Question 3

In which environmental compartments is photochemical transformation or photodegradation a potentially important degradation mechanism for organic chemicals and why is this the case? Explain with examples the differences between direct and indirect photodegradation in the environment.

3.7.2. Biodegradation

Author: John Parsons

Reviewers: Steven Droge, Russell Davenport

Leaning objectives:

You should be able to:

- the contribution of biochemical reactions in removing chemicals from the environment
- explain the differences between biotransformation, primary biodegradation and mineralization
- describe the most important biodegradation reactions under aerobic and anaerobic conditions

Keywords: Primary biodegradation, mineralisation, readily biodegradable chemicals, persistent chemicals, oxygenation reactions, reductions reactions

Introduction:

Biodegradation and biotransformation both refer to degradation reactions that are catalyzed by enzymes. In general, biodegradation is usually used to describe the degradation carried out by microorganisms and biotransformation often refers to reactions that follow the uptake of chemicals by higher organisms. This distinction is important and arises from the role that bacteria and other microorganisms play in natural biogeochemical cycles. As a result, microorganisms have the capacity to degrade most (perhaps all) naturally occurring organic chemicals in organic matter and convert them to inorganic end products. These reactions supply the microorganisms with the nutrients and energy they need to grow. This broad degradative capacity means that they are able to degrade many anthropogenic chemicals and potentially convert them to inorganic end products, a process referred to as mineralisation.

Although higher organisms are also able to degrade (metabolise) many anthropogenic chemicals, these chemicals are not taken up as source of nutrients and energy. Many anthropogenic chemicals can disturb cell functioning processes, and the biotransformation process has been proposed as a detoxification mechanism. Undesirable chemicals that may accumulate to potentially harmful levels are converted to products that are more rapidly excreted. In most cases, a polar and/or ionizable unit is attached to the chemical in one or two steps, making the compound more soluble in blood and more readily removed via the kidneys to the urine. This also renders most hazardous chemicals less toxic than the original chemical. Such biotransformation steps always costs energy (ATP, or through the use of e.g. NADH or NADPH in the enzymatic reactions) from the organism. Biotransformation is sometimes also used to describe degradation by microorganisms when this is limited to a conversion of a chemical into a new product.

Biodegradation is for many organic contaminants the major process that removes them from the environment. Measuring the rates of biodegradation therefore is a prominent aspect of chemical risk assessment. Internationally recognized standardised protocols have been developed to measure biodegradation rates of chemicals. Well know examples of these are the OCED Guidelines. These guidelines include screening tests designed to identify chemicals can be regarded as readily (i.e. rapidly) biodegradable as well as more complex tests to measure biodegradation rates of chemicals that degrade slowly in a variety of simulated environments. For more complex mechanistic studies, microorganisms able to degrade specific chemicals are isolated from environmental samples and cultivated in laboratory systems.

In principle, biodegradation of a chemical can be determined by either following the concentration of the chemical during the test or by following the conversion to end products (in most cases by either measuring oxygen consumption or CO_2 production). Although measuring the concentration gives the most directly relevant information on a chemical, it requires the availability or

development of analytical methods which is not always within the capability of routine testing laboratories. Measuring the conversion to CO_2 is comparatively straightforward but the production of CO_2 from other chemicals present in the test system (such as soil or dissolved organic matter) should be accounted for. This can be done by using ^{14}C-labelled chemicals in the tests but not all laboratories have facilities for this. The main advantage of this approach is that demonstration of quantitative conversion of a chemical to CO_2 etc. means that there is no concern about the accumulation of potentially toxic metabolites.

Since it is an enzymatically catalysed process, the rates of biodegradation should be modelled using the Michaelis Menten kinetics, or Monod kinetics if growth of the microorganisms is taken into account. In practice, however, first order kinetics are often used to model biodegradation in the absence of significant growth of the degrading microorganisms. This is more convenient that using Michaelis Menten kinetics but there is some justification for this simplification since the concentrations of chemicals in the environment are in general much lower than the half saturation concentrations of the degrading enzymes.

Table 1. Influence of molecular structure on the biodegradability of chemicals in the aerobic environment.

Type of compounds or substituents	**More biodegradable**	**Less biodegradable**
hydrocarbons	linear alkanes < C_{12}	linear alkanes > C_{12}
	alkanes with not too high molecular weight	high molecular weight alkanes
	linear chain	branched chain
	-C-C-C-	-C-O-C-
	aliphatic	aromatic
aliphatic chlorine	Cl more than 6 carbons from terminal C	Cl at less than 6 carbons from terminal C
Substituents to an aromatic ring	-OH	-F
	$-CO_2H$	-Cl
	$-NH_2$	$-NO_2$
	$-OCH_3$	$-CF_3$

Whether expressed as terms of first order kinetics or Michaelis Menten parameters, rates of biodegradation vary widely for different chemicals showing that chemical structure has a large impact on biodegradation. Large variations in biodegradation rates are however often observed for the same chemical in different experimental systems. This shows that environmental properties and conditions also play a key role in determining removal by biodegradation and it is often almost impossible to distinguish the effects of chemical properties from those of environmental properties. *In other words, there is no such thing as an intrinsic biodegradation rate of a chemical.* Nevertheless, we can derive some generic relationships between the structure and biodegradability of chemicals, as listed in Table 1. Examples are that branched hydrocarbon structures are degraded more slowly than linear hydrocarbon structures, and cyclic and in particular aromatic chemicals are degraded more slowly than aliphatic (non-aromatic) chemicals. Substituents and functional groups also have a major impact on biodegradability with halogens and other electron withdrawing substituents having strongly negative effects. It is therefore no surprise than the list of persistent organic pollutants is dominated by organohalogen compounds and in particular those with aromatic or alicyclic structures.

It should be recognized that biodegradation rates have often been observed to change over time. Long term exposure of microbial communities to new chemicals has often been observed to lead to increasing biodegradation rates. This phenomenon is called adaptation or acclimation and is often the case following repeated application of a pesticide at the same location. An example is shown for atrazine in Figure 2 where degradation rates increase following longer exposure to the pesticide.

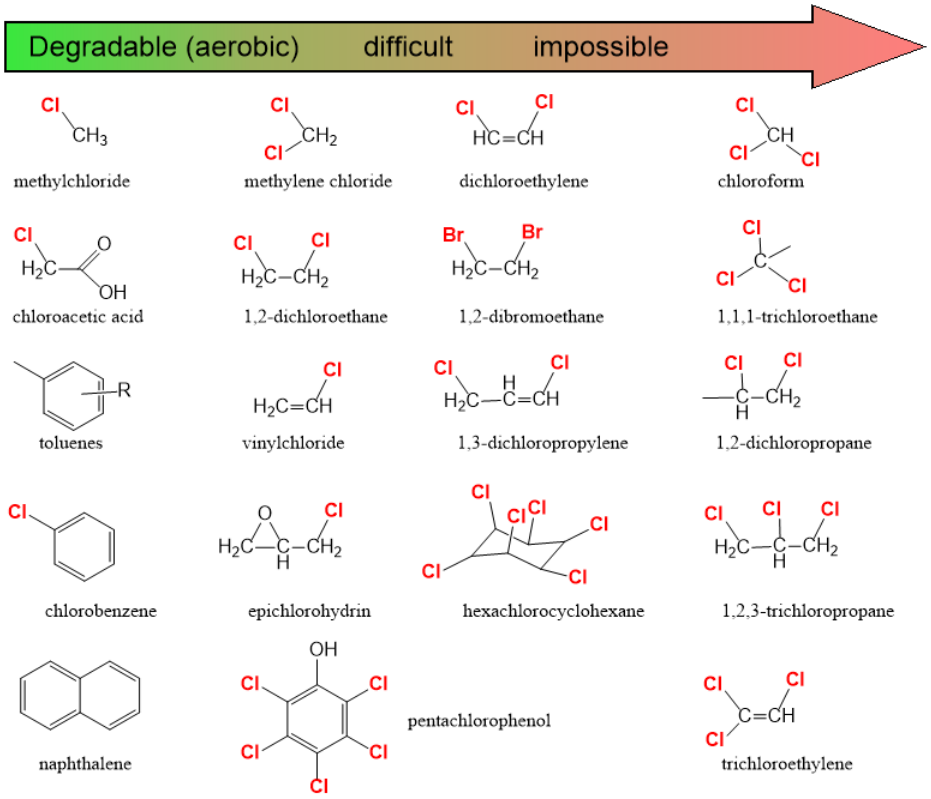


Figure 1. Effect of chlorination on (aerobic) biodegradation rates. Adapted from Janssen et al. (2005) by Steven Droge.

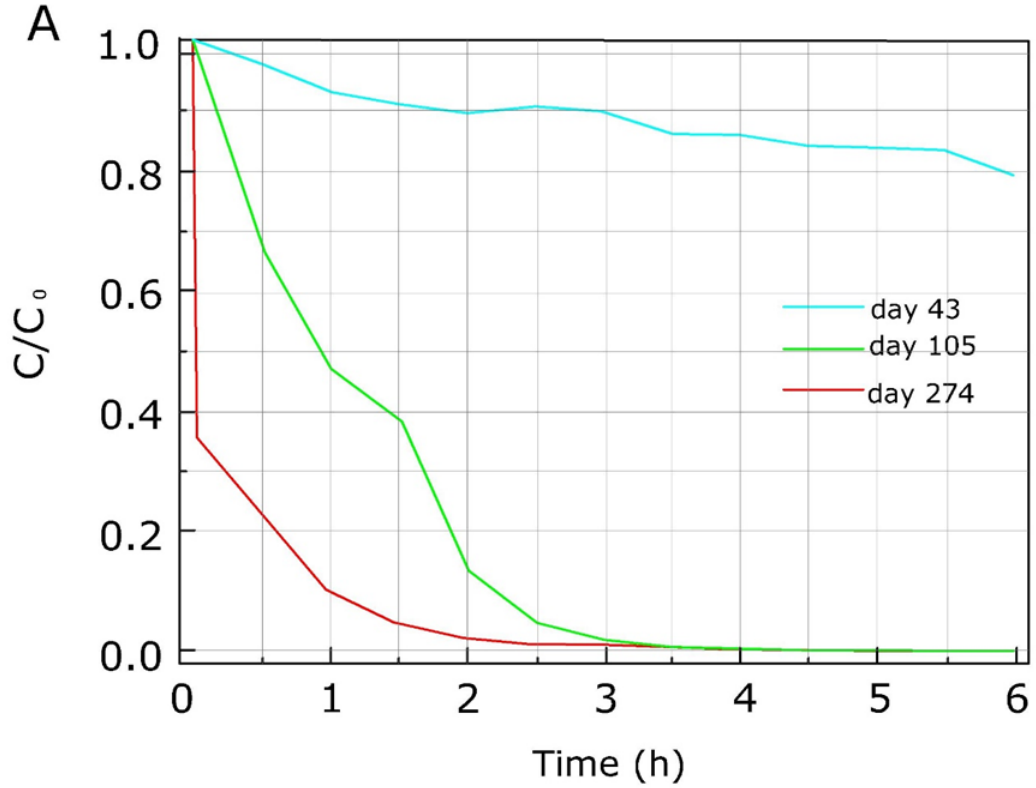


Figure 2. Comparison of the atrazine removal rates with (days 43 and 105) or without the addition of carbon and nitrogen sources (day 274). Redrawn from Zhou et al. (2017) by Wilma IJzerman.

Another recent example is the differences in biodegradation rates of the builder L-GLDA (tetrasodium glutamate diacetate) by activated sludge from different waste water treatment plants in the USA. Sludge from regions where L-GLDA was not or only

recently on the market required long lag time before degradation started whereas sludge from regions where L-GLDA -containing products had been available for several months required shorted lag phases.

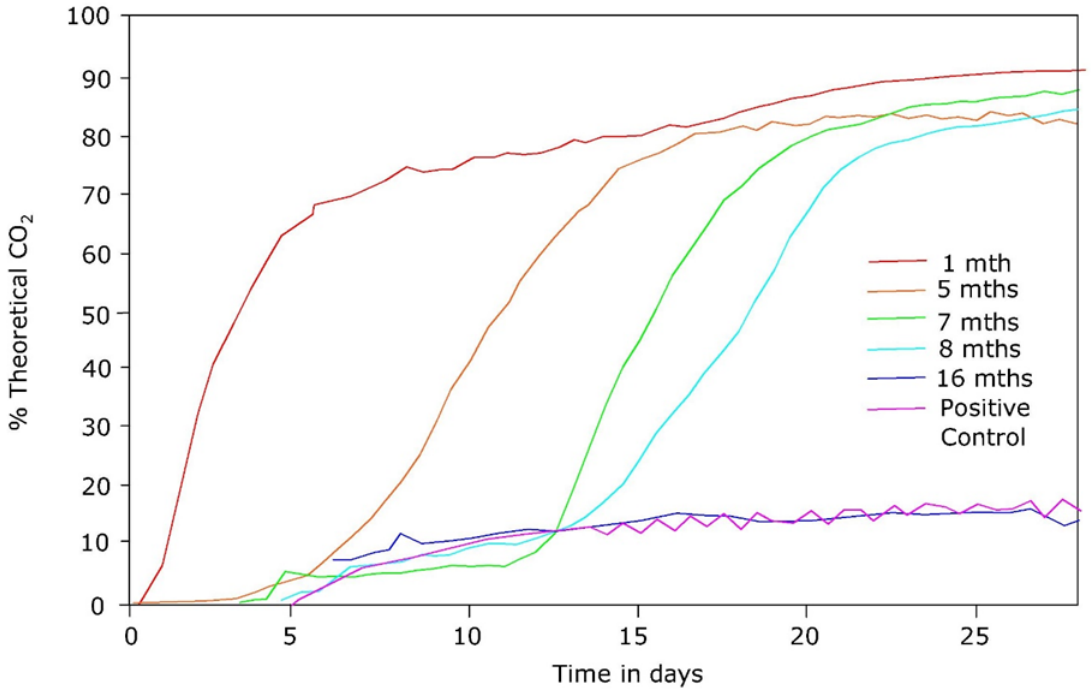


Figure 3. Biodegradation as a function of time following initial shipment of L-GLDA-containing products. Redrawn from Itrich et al. (2015) by Wilma Ijzerman.

Adaptation can results from i) shifts in composition or abundances of species in a bacterial community, ii) mutations within single populations, iii) horizontal transfer of DNA or iv) genetic recombination events, or combinations of these.

Biodegradation reactions and pathways

Biodegradation of chemicals that we regard as pollutants takes place when these chemicals are incorporated into the metabolism of microorganisms. The reactions involved in biodegradation are therefore similar to those involved in common metabolic reactions, such as hydrolyses, oxidations and reductions. Since the conversion of an organic chemical to CO_2 is an overall oxidation reaction, oxidation reactions involving molecular oxygen are probably the most important reactions. These reactions with oxygen are often the first but essential step in degradation and can be regarded as activation step converting relatively stable molecules to more reactive intermediates. This is particularly important for aromatic chemicals since oxygenation is required to make aromatic rings susceptible to ring cleavage and further degradation. These reactions are catalysed by enzymes called oxygenases of which there are broadly speaking two classes. Monoxygenases are enzymes catalysing reactions in which one oxygen atom of O_2 reacts with an organic molecule to produce a hydroxylated product. Examples of such enzymes are the cytochrome P450 family and are present in all organisms. These enzymes are for example involved in the oxidation of alkanes to carboxylic acids as part of the "beta-oxidation" pathway, which shortens linear alkanoic acids in steps of C_2-units, as shown in Figure 4.

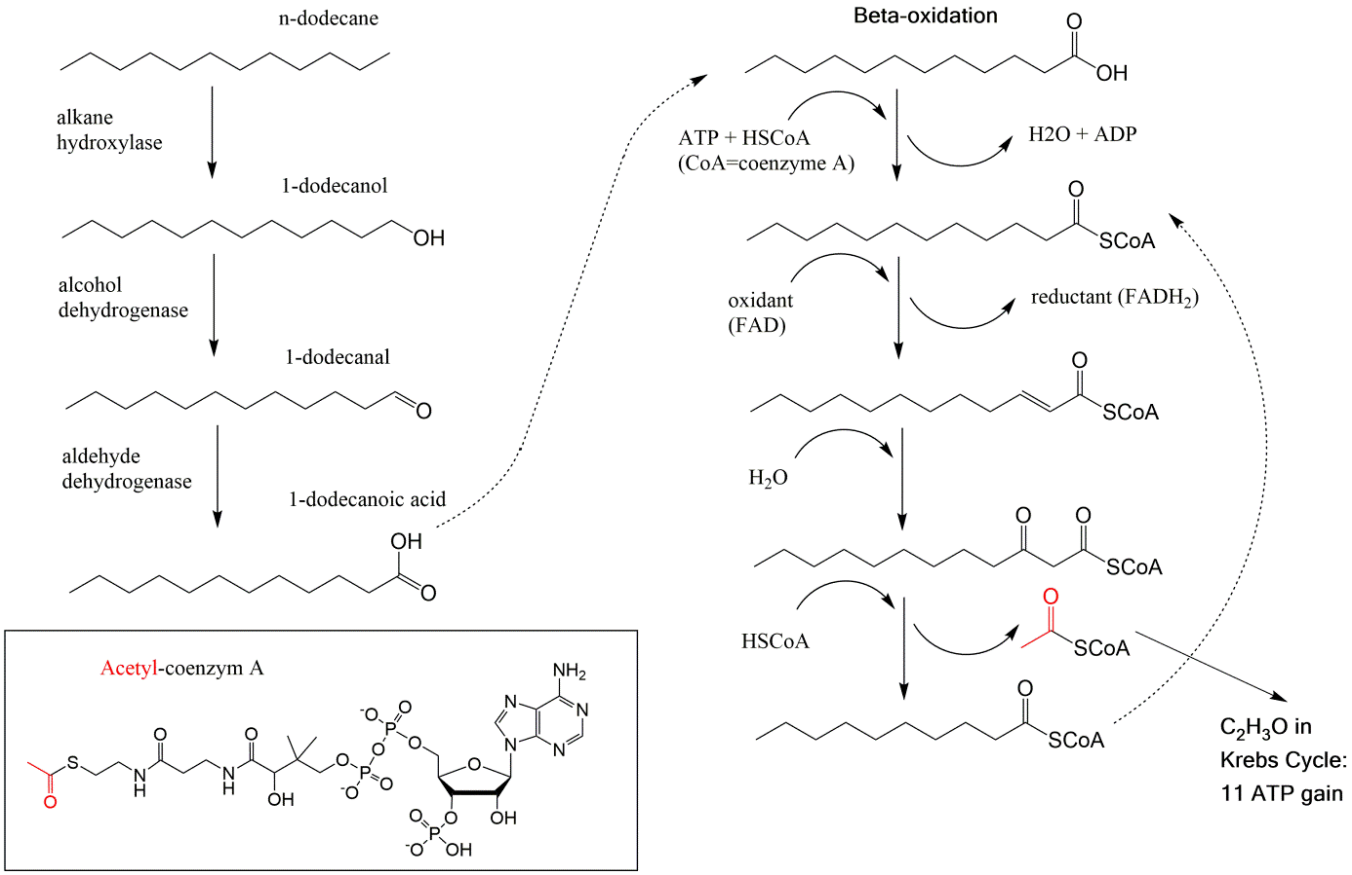


Figure 4. Typical oxidation steps of an alkane to an alkanoic acid, and subsequent beta-oxidation pathway from dodecanoic acid to decanoic acid, involving Coenzym-A. Redrawn from Schwarzenbach et al. (2003) by Steven Droge.

Dioxygenases are enzymes catalysing reactions in which both oxygen atoms of O_2 react with organic chemicals and appear to be unique to microorganisms such as bacteria. Examples of these reactions are shown for benzene in Figure 5. Similar reactions are involved in the degradation of more complex aromatic chemicals such as PAHs and halogenated aromatics.

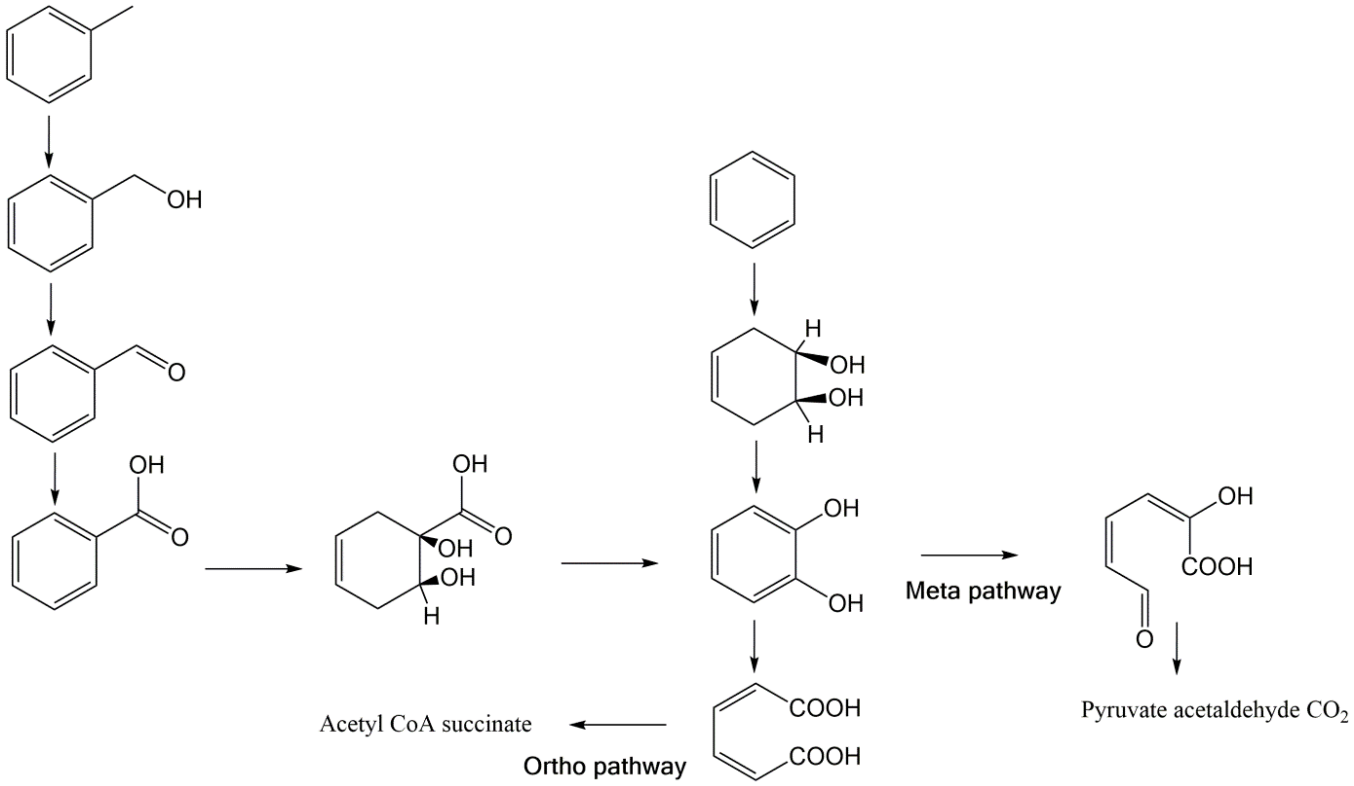


Figure 5. Examples of bacterial dioxygenation reactions and ring cleavage of toluene and benzene. Redrawn from Van Leeuwen and Vermeire (2007) by Steven Droge.

The absence of oxygen in anaerobic environments (sediments and groundwater) does not preclude oxidation of organic chemicals. Other oxidants present (nitrate, sulphate, Fe(III) etc) may be present in sufficiently high concentrations to act as oxidants and terminal electron acceptors supporting microbial growth. In the absence of oxygen, activation relies on other reactions, the most important reactions seem to be carboxylation or addition of fumarate. Figure 6 shows an example of the degradation of naphthalene to CO_2 in sediment microcosms under sulphate-reducing conditions.

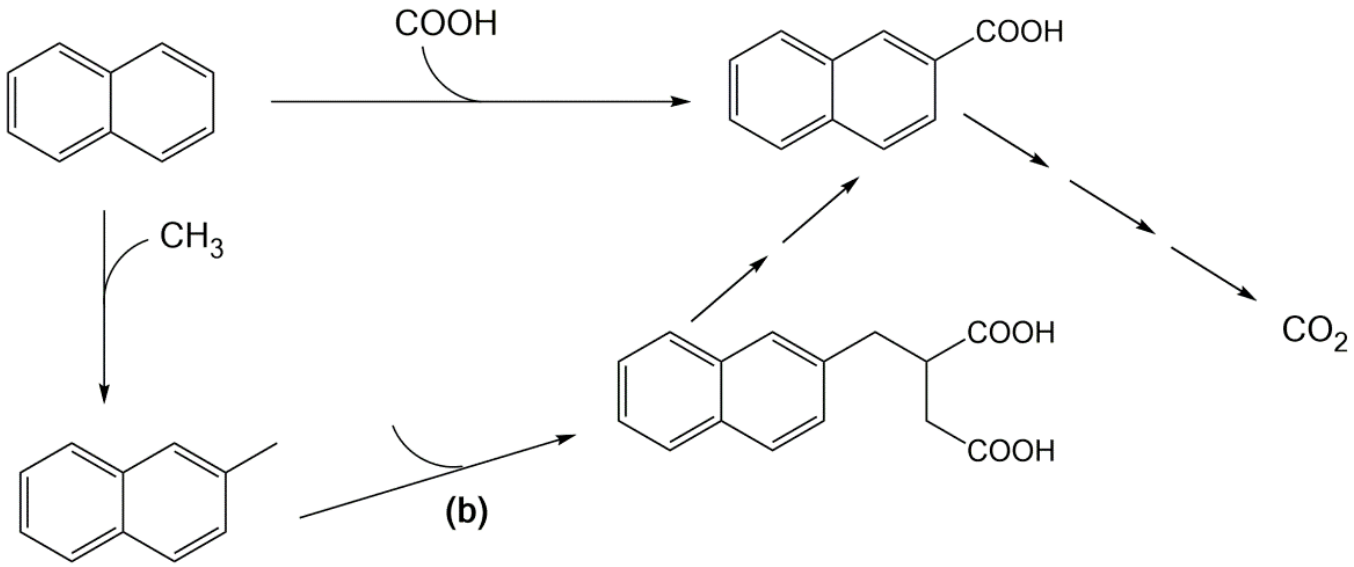


Figure 6. Proposed anaerobic oxidation pathway of naphthalene (redrawn from Kleeman and Merckenstock 2017). This involves two initial reaction mechanisms: carboxylation to naphthoic acid, methylation to 2-methylnaphthalene. Addition of fumarate (process b) could follow methylation and produce another way of transforming to 2-naphthoic acid. In both cases, 2-naphthoic acid is oxidized to CO_2.

Other important reactions in anaerobic ecosystems (sediments and groundwater plumes) are reductions. This affects functional groups, for example reduction of acids to aldehydes to alcohols, nitro groups to amino groups and, particularly important, substitution of halogens by hydrogen. The latter reactions can contribute to the conversion of highly chlorinated chemicals, that are resistant to oxidative biodegradation, to less chlorinated products which are more amenable to aerobic biodegradation. Many examples of these reductive dehalogenation reactions have been shown to occur in, for example, tetrachloroethene-contaminated groundwater (e.g. from dry-cleaning processes) and PCB-contaminated sediment. These reactions are exothermic under anaerobic conditions and some microorganisms are able to harvest this energy to support their growth. This can be considered to be a form of respiration based on dechlorination and is sometimes referred to as chlororespiration.

As is the case for abiotic degradation, hydrolyses are also important reactions in biodegradation pathways, particularly for chemicals that are derivatives of organic acids, such as carbamate, ester and organophosphate pesticides where hydrolyses are often the first step in their biodegradation. These reactions are similar to those described in the section on Chemical degradation.

References

Itrich, N.R., McDonough, K.M., van Ginkel, C.G., Bisinger, E.C., LePage, J.N., Schaefer, E.C., Menzies, J.Z., Casteel, K.D., Federle, T.W. (2015). Widespread microbial adaptation to L-glutamate-N,N,-diacetate (L-GLDA) following its market introduction in a consumer cleaning product. Environmental Science & Technology 49, 13314-13321.

Janssen, D. B., Dinkla, I. J. T., Poelarends, G. J., Terpstra, P. (2005). Bacterial degradation of xenobiotic compounds: evolution and distribution of novel enzyme activities, Environmental Microbiology 7, 1868-1882.

Kleemann, R., Meckenstock, R.U. (2017). Anaerobic naphthalene degradation by Gram-positive, iron-reducing bacteria. FEMS Microbial Ecology 78, 488-496.

Schwarzenbach, R.P., Gschwend, P.M., Imboden, D.M. (2017). Environmental Organic Chemistry, Third Edition, Wiley, ISBN 978-1-118-76723-8

Van Leeuwen, C., Vermeire, T.G. (2007). Risk Assessment of Chemicals: An Introduction (2nd ed.), Springer, ISBN 978-1-4020-6101-1

Zhou, Q., Chen, L. C., Wang, Z., Wang, J., Ni, S., Qiu, J., Liu, X., Zhang, X., Chen, X. (2017). Fast atrazine degradation by the mixed cultures enriched from activated sludge and analysis of their microbial community succession. Environmental Science & Pollution Research 24, 22152-22157.

3.7.2. Question 1

Degradation by microorganisms plays an important role in the environmental fate of industrial organic chemicals. Explain briefly why the role of microorganisms is so important.

Biodegradation rates depend on amongst other factors on the structure of the chemicals. Mention three structural factors responsible for slow biodegradation.

3.7.2. Question 2

DDT is one of the original "dirty dozen" Persistent Organic Pollutants (POPs). Explain what these POPs are and why they are labelled as persistent. What structural features are responsible for them being labelled as POPs?

3.7.2. Question 3

Groundwater used to prepare drinking water is discovered to be contaminated with toluene and tetrachloroethene. Although nothing is known about the geochemical conditions in the groundwater aquifer, you are asked to investigate whether there is any evidence for biodegradation of these compounds occurring in the aquifer. Suggest which compounds could be analysed as evidence for biodegradation.

3.7.3. Degradation test methods

Authors: John Parsons

Reviewers: Steven Droge, Russell Davenport

Leaning objectives:

You should be able to:

- explain the strategy used in standardised biodegradability testing
- describe the most important aspects of standard biodegradability testing protocols
- interpret the results of standardised biodegradability tests

Keywords: Environmental fate, chemical degradation, photochemical degradation, biodegradation, mineralisation, degradation rate

Introduction

Many experimental approaches are possible to measure the environmental degradation of chemicals, ranging from highly controlled laboratory experiments to environmental monitoring studies. While each of these approaches has its advantages and disadvantages, a standardised and relatively straightforward set of protocols has clear advantages such as suitability for a wide range of laboratories, broad scientific and regulatory acceptance and comparability for different chemicals.

The system of OECD test guidelines (see links in the reference list of this chapter) is the most important set of standardised protocols although other test systems may be used in other regulatory contexts. As well as tests covering environmental fate processes, they also cover physical-chemical properties, bioaccumulation, toxicity etc. These guidelines have been developed in an international context and are adopted officially after extensive validation and testing in different laboratories. This ensures their wide acceptance and application in different regulatory contexts for chemical hazard and risk assessment.

Chemical degradation tests

The OECD Guidelines include only two tests specific for chemical degradation. This might seem surprising but it should not be forgotten that chemical degradation could also contribute to the removal observed in biodegradability tests. The OECD Guidelines for chemical degradation are OECD Test 111: Hydrolysis as a Function of pH (OECD 2004) and OECD Test 316: Phototransformation of Chemicals in Water - Direct Photolysis (OECD 2008). If desired, sterilised controls may also be used to determine the contribution of chemical degradation in biodegradability tests.

OECD Test 111 measures hydrolytic transformations of chemicals in aquatic systems at pH values normally found in the environment (pH 4 - 9). Sterile aqueous buffer solutions of different pH values (pH 4, 7 and 9) containing radio-labelled or unlabelled test substance (below saturation) are incubated in the dark at constant temperature and analysed after appropriate time intervals for the test substance and for hydrolysis products. The preliminary test is carried out for 5 days at 50°C and pH 4.0, 7.0 and 9.0, this is known as a first tier test. Further second tier tests study the hydrolysis of unstable substances and the identification of hydrolysis products and may extend for 30 days.

OECD Test 316 measures direct photolysis rate constants using xenon arc lamp capable of simulating natural sunlight in the 290 to 800 nm or natural sunlight, and extrapolated to natural water. If estimated losses are superior or equal to 20%, the transformation pathway and the identities, concentrations, and rate of formation and decline of major transformation products are identified.

Biodegradability tests

Biodegradation is in general considered to be the most important removal process for organic chemicals in the environment and it is therefore no surprise that biodegradability testing plays a key role in the assessing the environmental fate and subsequent exposure risks of chemicals. Biodegradation is an extensively researched area but data from standardised tests are favoured for regulatory purposes as they are assumed to yield reproducible and comparable data. Standardised tests have been developed internationally, most importantly under the auspices of the OECD and are part of the wider range of tests to measure physical-chemical, environmental and toxicological properties of chemicals. An overview of these biodegradability tests is given in Table 1.

The way that biodegradability testing is implemented can vary in detail depending on the regulatory context but in general it is based on a tiered approach with all chemicals being subjected to screening tests to identify chemicals that can be considered to be readily biodegradable and therefore removed rapidly from wastewater treatment plants (WWTPs) and the environment in general. These tests were originally developed for surfactants and often use activated sludge from WWTPs as a source of microorganisms since biodegradation during wastewater treatment is a major conduit of chemical emissions to the environment. The so-called ready biodegradability tests are designed to be stringent with low bacterial concentrations and the test chemical as the only potential

source of carbon and energy at high concentrations. The assumption is that chemicals that show rapid biodegradation under these unfavourable conditions will always be degraded rapidly under environmental conditions. Biodegradation is determined as conversion to CO_2 (mineralisation), either by directly measuring CO_2 produced, or the consumption of oxygen, or removal of dissolved organic carbon, as this is the most desirable outcome of biodegradation. The results that have to be achieved for a chemical to be considered readily biodegradable vary slightly depending on the test, but as an example in the OECD 301D test (OECD 2014), the consumption of oxygen should reach 70% of that theoretically required for complete mineralisation within 28 days.

Table 1. The OECD biodegradability tests

OECD TEST GUIDELINE	PARAMETER MEASURED	REFERENCE
Ready biodegradability tests		
301A: DOC Die-away test	DOC	OECD 1992a
301B: CO_2 evolution test	CO_2	OECD 1992a
301C: Modified MITI(I) test	O_2	OECD 1992a
301D: Closed bottle test	O_2	OECD 1992a
301E: Modified OECD screening test	DOC	OECD 1992a
301F: Manometric respirometry test	O_2	OECD 1992a
306: Biodegradability in seawater	DOC	OECD 1992c
310: Test No. 310: Ready Biodegradability - CO_2 in sealed vessels (Headspace Test).	CO_2	OECD 2014
Inherent biodegradability tests		
302A: Modified Semi-continuous Activated Sludge (SCAS) test	DOC	OECD 1981b
302B: Zahn-Wellens test	DOC	OECD 1992b
302C: Modified MITI(II) test	O_2	OECD 2009
Simulation tests		
303A: Activated sludge units	DOC	OECD 2001
303B: Biofilms	DOC	OECD 2001
304A: Inherent biodegradability in soil	$^{14}CO_2$	OECD 1981a
307: Aerobic and anaerobic transformation in soil	$^{14}CO_2/CO_2$	OECD 2002a
308: Aerobic and anaerobic transformation in aquatic sediment systems	$^{14}CO_2/CO_2$	OECD 2002b
309: Aerobic mineralization in surface water	$^{14}CO_2/CO_2$	OECD 2004b
311: Anaerobic biodegradability of organic compounds in digested sludge: by measurement of gas production	CO_2 and CH_4	OECD 2006
314: Simulation tests to assess the biodegradability of chemicals discharged in wastewater	Concentration of chemical, $^{14}CO_2/CO_2$	OECD 2008a

These test systems are widely applied for regulatory purposes but they do have a number of issues. These include the fact that there are practical difficulties when applied to volatile or poorly soluble chemicals, but probably the most important is that for some chemicals the results can be highly variable. This is usually attributed to the source of the microorganisms used to inoculate the system. For many chemicals, there is a wide variability in how quickly they are degraded by activated sludge from different WWTPs. This is probably the result of different exposure concentrations and exposure periods to the chemicals, and may also be caused by dependence on the ability of small populations of degrading microorganisms, which may not always be included in the sludge samples used in the tests. These issues are not dealt with in any systematic way in biodegradability testing. It has been suggested that a preliminary period of exposure to the chemicals to be tested would allow sludge to adapt to the chemicals and may yield more reproducible test results. Further suggestions include using a higher, more environmentally relevant, concentration of activated sludge as the inoculum.

Failure to comply with the pass criteria in ready biodegradability tests does not necessarily mean that the chemical is persistent in the environment since it is possible that slow biodegradation may occur. These chemicals may therefore be tested further in higher tier tests, for what is referred to as inherent biodegradability in tests performed under more favourable conditions or in simulation tests representing specific compartments, to determine whether biodegradation may contribute significantly to their removal. These tests are also standardised (see Table 1). Simulation tests are designed to represent environmental conditions in specific compartments, such as redox potential, pH, temperature, microbial community, concentration of test substance and occurrence and concentration of other substrates.

The criteria used in classifying the biodegradability of chemicals depend on the regulatory context. Biodegradability tests can be used for different purposes: in the EU this includes 3 distinct purposes; classification and labelling, hazard/persistent assessment, and environmental risk assessment Recently regulatory emphasis has shifted to identifying hazardous chemicals, and therefore those chemicals that are less biodegradable and likely to persist in the environment. Examples for the classification as PBT (persistent, bioaccumulative and toxic) or vPvB (very persistent and very bioaccumulative) chemicals are shown in Table 2. As well as the results of standardised tests, other data such as the results of environmental monitoring data or studies on the microbiology of biodegradation can also be taken into account in evaluations of environmental degradation in a so-called weight of evidence approach.

Table 2. Criteria used to classify chemicals as PBT or vPvB (van Leeuwen & Vermeire 2007)

Property	**PBT criteria**	**vPvB criteria**
Persistence	$T_{1/2}$ >60 days in marine water, or $T_{1/2}$ >40 days in fresh/estuarine water, or $T_{1/2}$ >180 days in marine sediment, or $T_{1/2}$ >120 days in fresh/estuarine sediment, or $T_{1/2}$ >120 days in soil.	$T_{1/2}$ >60 days in marine, fresh or estuarine water, or $T_{1/2}$ >180 days in marine, fresh or estuarine sediment, or $T_{1/2}$ >180 days in soil
Bioaccumulation	BCF > 2000 L/kg	BCF > 5000 L/kg
Toxicity	- NOEC < 0.01 mg/L for marine or freshwater organisms, or - substance is classified as carcinogenic, mutagenic, or toxic for reproduction, or - there is other evidence of chronic toxicity, according to Directive 67/548/EEC	

The results of biodegradability tests are sometimes also used to derive input data for environmental fate models (see section on Multicompartment modeling). It is however not always straightforward to transfer data measured in what is sometimes a multi-compartment test system into degradation rates in individual compartments as other processes (e.g. partitioning) need to be taken into account.

References

OECD, 1981a. OECD Guidelines for the Testing of Chemicals. Test No. 304A: Inherent Biodegradability in Soil.

OECD, 1981b. OECD Guidelines for the Testing of Chemicals. Test No. 302A: Inherent Biodegradability: Modified SCAS Test.

OECD, 1992a. OECD Guidelines for the Testing of Chemicals. Test No. 301: Ready Biodegradability.

OECD, 1992b. OECD Guidelines for the Testing of Chemicals. Test No. 302B: Inherent Biodegradability: Zahn-Wellens/ EVPA Test.

OECD, 1992c. OECD Guidelines for the Testing of Chemicals. Test No. 306: Biodegradability in Seawater.

OECD, 2001. OECD Guidelines for the Testing of Chemicals. Test No. 303: Simulation Test - Aerobic Sewage Treatment - A: Activated Sludge Units; B: Biofilms.

OECD, 2002a. OECD Guidelines for the Testing of Chemicals. Test No. 307: Aerobic and Anaerobic Transformation in Soil.

OECD, 2002b. OECD Guidelines for the Testing of Chemicals. Test No. 308: Aerobic and Anaerobic Transformation in Aquatic Sediment Systems.

OECD, 2004a. OECD Guidelines for the Testing of Chemicals. Test No. 111: Hydrolysis as a Function of pH.

OECD, 2004b. OECD Guidelines for the Testing of Chemicals. Test No. 309: Aerobic Mineralisation in Surface Water - Simulation Biodegradation Test

OECD, 2006. OECD Guidelines for the Testing of Chemicals. Test No. 311: Anaerobic Biodegradability of Organic Compounds in Digested Sludge: by Measurement of Gas Production.

OECD, 2008a. OECD Guidelines for the Testing of Chemicals. Test No. 314: Simulation Tests to Assess the Biodegradability of Chemicals Discharged in Wastewater

OECD, 2008b. OECD Guidelines for the Testing of Chemicals. Test No. 316: Phototransformation of Chemicals in Water - Direct Photolysis.

OECD, 2009. OECD Guidelines for the Testing of Chemicals. Test No. 302C: Inherent Biodegradability: Modified MITI Test (II).

OECD, 2014. OECD Guidelines for the Testing of Chemicals. Test No. 310: Ready Biodegradability - CO_2 in sealed vessels (Headspace Test).

Van Leeuwen, C.J., Vermeire, T.G. (2007). Risk Assessment of Chemicals: An Introduction (2nd ed.), Springer, ISBN 978-1-4020-6101-1

3.7.3. Question 1

The OECD has published extensive protocols that can be used to evaluate the degradability of chemicals in the environment. Which forms of abiotic (chemical) degradation do these include?

3.7.3. Question 2

Biodegradation tests used to evaluate the environmental effects of chemicals often measure the mineralisation of the chemicals, for example by measuring the amount of carbon dioxide produced. Explain why this is preferred to measuring the removal of the original chemical (primary biodegradation).

3.7.3. Question 3

Assessment of the biodegradability of chemicals often follows a tiered system in which they are first screened for their ready biodegradability before undergoing more extensive testing to assess their inherent biodegradability or biodegradation in specific environmental compartments using simulation tests. Explain why this approach is applied.

This page titled 2.3.4.6: Degradation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.6: Degradation** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.4.7: Modelling exposure

3.8. Modelling exposure

3.8.1. How to estimate emission?

In preparation

3.8.2. Multicompartment modeling

Authors: Dik van de Meent and Michael Matthies

Reviewer: John Parsons

Learning objectives:

You should be able to

- explain what a mass balance equation is
- describe how mass balance equations are used in multimedia fate modeling
- explain the concepts of thermodynamic equilibrium and steady state
- give some examples of the use of multimedia mass balance modeling

Keywords: mass balance equation, environmental fate model

The mass balance equation

Multicompartment (or multimedia) mass balance modeling starts from the universal conservation principle, formulated as balance equation. The governing principle is that the rate of change (of any entity, in any system) equals the difference between the sum of all inputs (of that entity) to the system and the sum of all outputs from it. Environmental modelers use the balance equation to predict exposure concentrations of chemicals in the environment by deduction from knowledge of the rates of input- and output processes, which can be understood easiest from considering the mass balance equation for one single environmental compartment (Figure 1):

$$\frac{dm^{i,j}}{dt} = \sum^{j} input^{i,j} - \sum^{j} output^{i,j} \quad \text{(eq. 1)}$$

where $dm^{i,j}/dt$ represents the change of mass of chemical i in compartment j (kg) over time (s), and $input^{i,j}$ and $output^{i,j}$ denote the rates of input and output of chemical to and from compartment j, respectively.

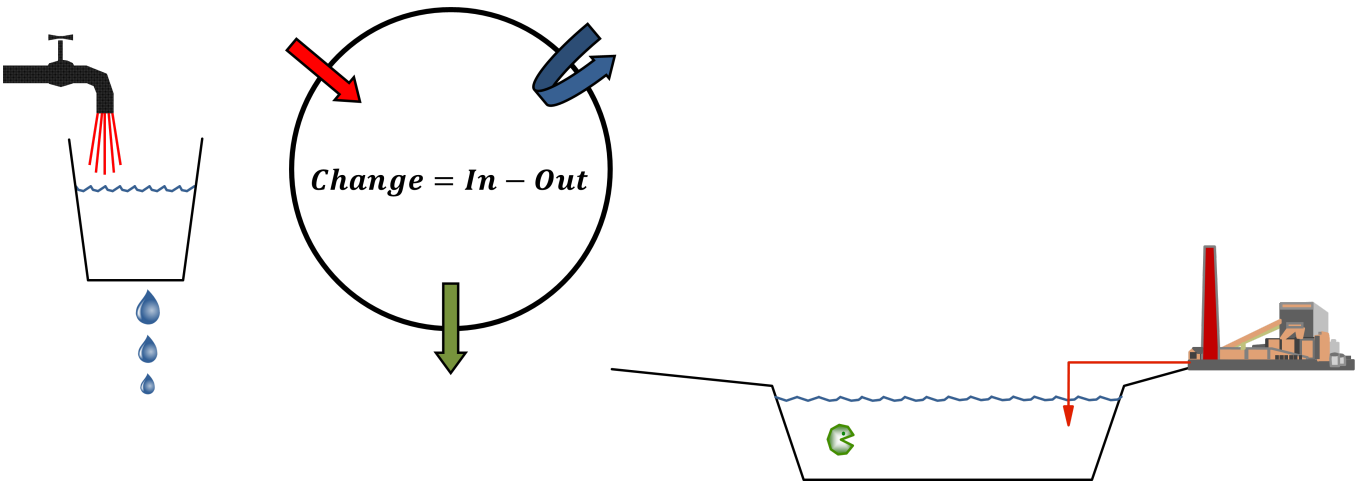


Figure 1. The mass of a chemical in a lake is like the mass of water in a leaking bucket: both can be described with the universal (mass) balance equation, which says differences in inputs and outputs make amounts held up in systems change: $\frac{dm}{dt} = \sum inputs - \sum outputs$.

One compartment model

In multimedia mass balance modeling, mass balance equations (of the type shown in equation 1) are formulated for each environmental compartment. Outflows of chemical from the compartments are often proportional to the amounts of chemical present in the compartments, while external inputs (emissions) may often be assumed constant. In such cases, i.e. when first-order kinetics apply (see section 3.3 on Environmental fate of chemicals), mass balance equations take the form of equation 1 in section 3.3. For one compartment (e.g. a lake, as in Figure 1) only:

$\frac{dm}{dt} = I - k\,m$ (eq. 2)

in which d*m*/d*t* (kg.s^{-1}) is the rate of change of the mass (kg) of chemical in the lake, *I* (kg.s^{-1}) is the (constant) emission rate, and the product *k.m* (kg.s^{-1}) denotes the first-order loss rate of the chemical from the lake. It is obvious that eventually a steady state must develop, in which the mass of chemical in the lake reaches a predictable maximum

$\frac{dm}{dt} = I - k\,m = 0 \rightarrow m_\infty = \frac{I}{k}$ (eq. 2a)Figure 2):

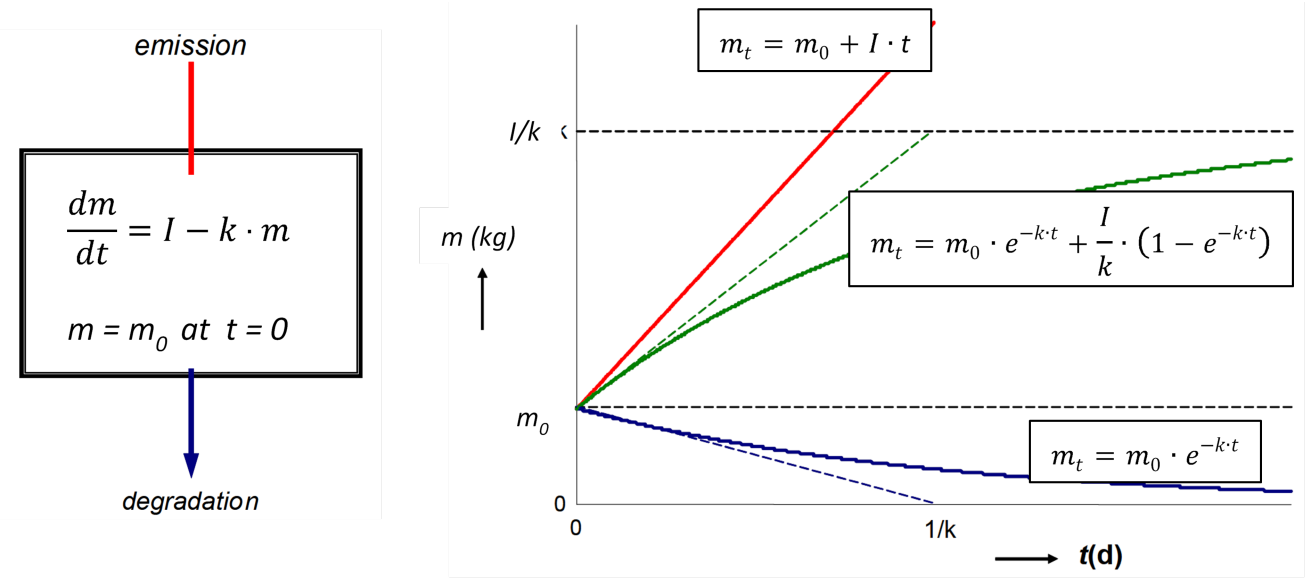


Figure 2. For one compartment only, in case loss processes obey first-order kinetics and emissions are constant (i.e. not varying with time), the mass of chemical is expected to increase exponentially from its initial mass m_0 to its steady-state level I, which will be reached at ininite time t_∞. *After Van de Meent et al. (2011).*

$m_t = m_0\, e^{-k\,t} + \frac{I}{k}\,(1 - e^{-k\,t}),\ with\ \frac{I}{k} = m_\infty$ (eq. 3)

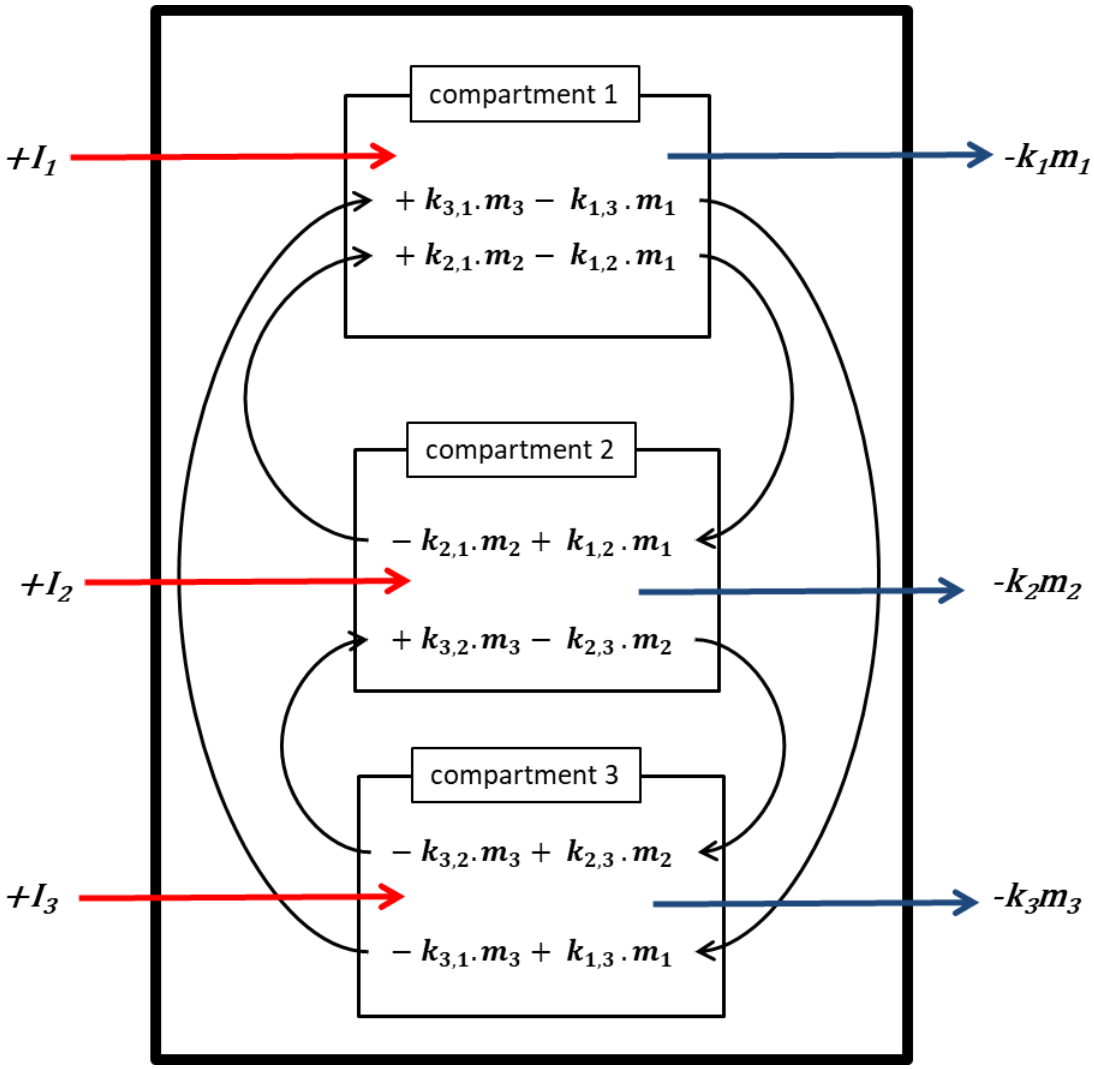
When the input rate (emission) is constant, i.e. that it does not vary with time, and is independent of the mass of chemical present, the mass of chemical in the systems is expected to increase exponentially, from its initial value at



t_0, to a steady level at t_∞. According to equation 3, a final mass level equal to $\frac{I}{k}$ is to be expected.

Multi-compartment model

The prefix 'multi' indicates that generally (many) more than one environmental compartment is considered. The Unit World (see below) contains air, water, biota, sediment and soil; more advanced global modeling systems may use hundreds of compartments. The case of three compartments (typically one air, one water, one soil) is schematically worked out in Figure 3.



$$\frac{dm_1}{dt} = I_1 - (k_1 + k_{1,2} + k_{1,3})\, m_1 + k_{2,1}\, m_2 + k_{3,1}\, m_3$$

$$\frac{dm_2}{dt} = I_2 + k_{1,2}\, m_1 - (k_2 + k_{2,1} + k_{2,3})\, m_2 + k_{3,2}\, m_3$$

$$\frac{dm_3}{dt} = I + k_{1,3}\, m_1 + k_{2,3}\, m_2 - (k_3 + k_{3,1} + k_{3,2})\, .\, m_3$$

Figure 3. Three-compartment mass balance model. Arrows represent mass flows of chemical to and from compartments. Losses from source compartments (negative signs) become gains to receiving compartments (positive sign). The model consists of three differential mass balance equations, with (nine) known rate constants $k_{i,j}$ (for flows out of source compartments i, into receiving compartments j, in s^{-1}), and (three) unknown masses m_i (kg). From Van de Meent et al. (2011), with permission.

Each compartment can receive constant inputs (emissions, imports), and chemical can be exported from each compartment by degradation or advective outflow, as in the one-compartment model. In addition, chemical can be transported between compartments (simultaneous import-export). All mass flows are characterized by (pseudo) first-order rate constants (see section 3.3 on Environmental fate processes). The three mass balance equations eventually balance to zero at infinite time:

$$balance_1 = 0 = I_1 - (k_1 + k_{1,2} + k_{1,3})\, m_1^* + k_{2,1}\, m_2^* + k_{3,1}\, m_3^*$$

$$balance_2 = 0 = I_2 + k_{1,2}\, m_1^* - (k_2 + k_{2,1} + k_{2,3})\, m_2^* + k_{3,2}\, m_3^*$$

$$balance_3 = 0 = I_3 + k_{1,3}\, m_1^* + k_{2,3}\, m_2^* - (k_3 + k_{3,1} + k_{3,2})\, m_3^* \quad \text{(eq. 4)}$$

where the symbols m_i^* denote mass in compartments i at steady state. Sets of n linear equations with n unknowns can be solved algebraically, by manually manipulating equations 4, until clean expressions for each of the three m_i values are obtained, which

inevitably becomes tedious as soon as more than two mass balance equations are to be solved - this did not withhold one of Prof. Mackay's most famous PhD students from successfully solving a set of 14 equations! An easier way of solving sets of n linear equations with n unknowns is by means of linear algebra. Using linear algebraic vector-matrix calculus, the equations 4 can be rewritten into one linear-algebraic equation:

$$\frac{dm_{\bigcirc}^{\rightarrow}}{dt} = 0 = I_{\bigcirc}^{\rightarrow} + A\, m_{\bigcirc}^{\rightarrow} \quad \text{(eq. 5)}$$

in which in which $m_{\bigcirc}^{\rightarrow}$ is the vector of masses in the three compartments, A is the model matrix of known rate constants and $e_{\bigcirc}^{\rightarrow}$ is the vector of known emission rates:

$$m_{\bigcirc}^{\rightarrow} = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix},\ I_{\bigcirc}^{\rightarrow} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix},\ A = \begin{bmatrix} -(k_1 + k_{1,2} + k_{1,3}) & k_{2,1} & k_{3,1} \\ k_{1,2} & -(k_2 + k_{2,1} + k_{2,3}) & k_{3,2} \\ k_{1,3} & k_{2,3} & -(k_3 + k_{3,1} + k_{3,2}) \end{bmatrix}$$

The solution of equation 5 is

$$m_{\bigcirc}^{s} = -A^{-1}\, I_{\bigcirc}$$

in which $m_{\bigcirc}^{s}$ is the vector of masses at steady state and A^{-1} is the inverse of model matrix A. The linear algebraic method of solving linear mass balance equations is easily carried with spreadsheet software (such as MS Excel, LibreOffice Calc or Google Sheets), which contain built-in array functions for inverting matrices and multiplying them by vectors.

Unit World modeling

In the late 1970s, pioneering environmental scientists at the USEPA Environmental Research Laboratory in Athens GA, recognized that the universal (mass) balance equation, applied to compartments of environmental media (air, water, biota, sediment, soil) could serve as a means to analyze and understand differences in environmental behavior and fate of chemicals. Their 'evaluative Unit World Modeling' (Baughman and Lassiter, 1978; Neely and Blau, 1985) was the start of what is now known as multimedia mass balance modeling. The Unit World concept was further developed and polished by Mackay and co-workers (Neely and Mackay, 1982; Mackay and Paterson, 1982; Mackay et al., 1985; Paterson and Mackay, 1985, 1989). In Unit World modeling, the environment is viewed of as a set of well-mixed chemical reactors, each representing one environmental medium (compartment), to and from which chemical flows, driven by 'departure from equilibrium' - this is chemical technology jargon for expressing the degree to which thermodynamic equilibrium properties such as 'chemical potential' or 'fugacity' differ (Figure 4). Mackay and co-workers used fugacity in mass balance modeling as the central state variable. Soon after publication of this 'fugacity approach' (Mackay, 1991), the term 'fugacity model' became widely used to name all models of the 'Mackay-type', which applied 'Unit World mass balance modeling', even though most of these models kept using the more traditional chemical mass as a state variable.

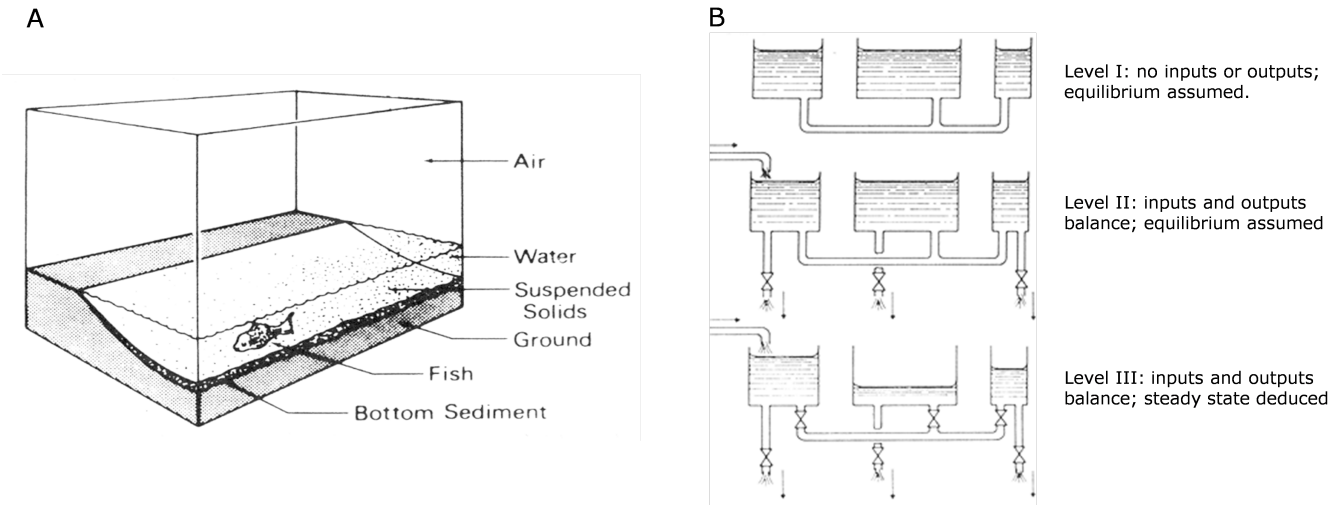


Figure redrawn after Mackay (1991).

Complexity levels

While conceptually simple (environmental fate is like a leaking bucket, in the sense that its steady-state water height is predictable from first-order kinetics), the dynamic character of mass balance modeling is often not so intuitive. The abstract mathematical perspective may best suit explain mass balance modeling, but this may not be practical for all students. In his book about multimedia mass balance modeling, Mackay chose to teach his students the intuitive approach, by means of his famous water tank analogy (Figure 4B).

According to this intuitive approach, mass balance modeling can be done at levels of increasing complexity, where the lowest, simplest, level that serves the purpose should be regarded as the most suitable. The least complex is level I assuming no input and output. A chemical can freely (i.e. without restriction) flow from one environmental compartment to another, until it reaches its state of lowest energy: the state of thermodynamic equilibrium. In this state, the chemical has equal chemical potential and fugacity in all environmental media. The system is at rest; in the hydraulic analogy, water has equal levels in all tanks. This is the lowest level of model complexity, because this model only requires knowledge of a few thermodynamic equilibrium constants, which can be reasoned from basic physical substance properties.

The more complex modeling level III describes an environment in which flow of chemical between compartments experiences flow resistance, so that a steady state of balance between outputs and inputs is reached only at the cost of permanent 'departure from equilibrium'. Degradation in all compartments and advective flows, e.g. rain fall or wind and water currents, are also considered. The steady state of level III is one in which fugacities of chemical in the compartments are unequal (no thermodynamic equilibrium); in the hydraulic analogy, water in the tanks rest at different heights. Naturally, solving modeling level III requires detailed knowledge of the inputs (into which compartment(s) is the chemical emitted?), the outputs (at what rates is the chemical degraded in the various compartments?) and the transfer resistances (how rapid or slow is the mass transfer between the various compartments?). Level III modelers are rewarded for this by obtaining more realistic model results.

The fourth complex level of multimedia mass balance modeling (level IV, not shown in Figure 4B) produces transient (time dependent) solutions. Model simulations start ($t = 0$) with zero chemical ($m = 0$; empty water tanks). Compartments (tanks), fill up gradually until the system comes to a steady state, in which generally one or more compartments depart from equilibrium, as in level III modeling. Level IV is the most realistic representation of environmental fate of chemicals, but requires most detailed knowledge of mass flows and mass transfer resistances. Moreover, time-varying states are least easy to interpret and not always most informative of chemical fate. The most important piece of information to be gained from level IV modeling is the indication of time to steady state: how long does it take to clear the environment from persistent chemicals that are no longer used?

Mackay describes an intermediate level of complexity (level II), in which outputs (degradation, advective outflows) balance inputs (as in level III), and chemical is allowed to freely flow between compartments (as in level I). A steady state develops in level II and there is thermodynamic equilibrium at all times. Modeling at level II does not require knowledge of mass transfer resistances (other than that resistances are negligible!), but degradation and outflow rates increase the model complexity compared to that of level I. In many situations, level II modeling yields surprisingly realistic results.

Use of multimedia mass balance models

Soon after publication of the first use of 'evaluative Unit World modeling' (Mackay and Paterson, 1982), specific applications of the 'Mackay approach' to multimedia mass balance modeling started to appear. The Mackay group published several models for the evaluation of chemicals in Canada, of which ChemCAN (Mackay et al., 1995) is known best. Even before ChemCAN, the Californian model CalTOX (Mckone, 1993) and the Dutch model SimpleBox (Van de Meent, 1993) came out, followed by publication of the model HAZCHEM by the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC, 1994) and the German Umwelt Bundesamt's model ELPOS (Beyer and Matthies, 2002). Essentially, all these models serve the very same purpose as the original Unit World model, namely providing standardized modeling platforms for evaluating the possible environmental risks from societal use of chemical substances.

Multimedia mass balance models became essential tools in regulatory environmental decision making about chemical substances. In Europe, chemical substances can be registered for marketing under the REACH regulation only when it is demonstrated that the chemical can be used safely. Multimedia mass balance modeling with SimpleBox (Hollander et al., 2014) and SimpleTreat (Struijs et al, 2016) plays an important role in registration.

While early multimedia mass balance models all followed in the footsteps of Mackay's Unit World concept (taking the steady-state approach and using one compartment per environmental medium), later models became larger and spatially and temporally explicit, and were used for in-depth analysis of chemical fate.

In the late 1990s, Wania and co-workers developed a Global Distribution Model for Persistent Organic Pollutants (GloboPOP). They used their global multimedia mass balance model to explore the so-called cold condensation effect, by which they explained the occurrence of relatively large amounts of persistent organic chemicals in the Arctic, where no one had ever used them (Wania, 1999). Scheringer and co-workers used their CliMoChem model to investigate long-range transport of persistent chemicals into Alpine regions (Scheringer, 1996; Wegmann et al., 2005). MacLeod and co-workers (Toose et al., 2004) constructed a global multimedia mass balance model (BETR World) to study long-range, global transport of pollutants.

References

Baughman, G.L., Lassiter, R. (1978). Predictions of environmental pollutant concentrations. In: Estimating the Hazard of Chemical Substances to Aquatic Life. ASTM STP 657, pp. 35-54.

Beyer, A., Matthies, M. (2002). Criteria for Atmospheric Long-range Transport Potential and Persistence of Pesticides and Industrial Chemicals. Umweltbundesamt Berichte 7/2002, E. Schmidt-Verlag, Berlin. ISBN 3-503-06685-3.

ECETOC (1994). HAZCHEM, A mathematical Model for Use in Risk Assessment of Substances. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels.

Hollander, A., Schoorl, M., Van de Meent, D. (2016). SimpleBox 4.0: Improving the model, while keeping it simple... Chemosphere 148, 99-107.

Mackay, D. (1991). Multimedia Environmental Fate Models: The Fugacity Approach. Lewis Publishers, Chelsea, MI.

Mackay, D., Paterson, S. (1982). Calculating fugacity. Environmental Science and Technology 16, 274-278.

Mackay, D., Paterson, S., Cheung, B., Neely, W.B. (1985). Evaluating the environmental behaviour of chemicals with a level III fugacity model. Chemosphere 14, 335-374.

Mackay, D., Paterson, S., Tam, D.D., Di Guardo, A., Kane, D. (1995). ChemCAN: A regional Level III fugacity model for assessing chemical fate in Canada. Environmental Toxicology and Chemistry 15, 1638-1648.

McKone, T.E. (1993). CALTOX, A Multi-media Total-Exposure Model for Hazardous Waste sites. Lawrence Livermore National Laboratory. Livermore, CA.

Neely, W.B., Blau, G.E. (1985). Introduction to Exposure from Chemicals. In: Neely, W.B., Blau, G.E. (Eds). Environmental Exposure from Chemicals Volume I, CRC Press, Boca Raton, FL., pp 1-10.

Neely, W.B., Mackay, D. (1982). Evaluative model for estimating environmental fate. In: Modeling the Fate of Chemicals in the Aquatic Environment. Ann Arbor Science, Ann Arbor, MI, pp. 127-144.

Paterson, S. (1985). Equilibrium models for the initial integration of physical and chemical properties. In: Neely, W.B., Blau, G.E. (Eds). Environmental Exposure from Chemicals Volume I, CRC Press, Boca Raton, FL., pp 218-231.

Paterson, S., Mackay, D. (1989). A model illustrating the environmental fate, exposure and human uptake of persistent organic chemicals. Ecological Modelling 47, 85-114.

Scheringer, M. (1996). Persistence and spatial range as endpoints of an exposure-based assessment of organic chemicals. Environmental Science and Technology 30, 1652-1659.

Struijs, J., Van de Meent, D., Schowanek, D., Buchholz, H., Patoux, R., Wolf, T., Austin, T., Tolls, J., Van Leeuwen, K., Galay-Burgos, M. (2016). Adapting SimpleTreat for simulating behaviour of chemical substances during industrial sewage treatment. Chemosphere 159:619-627.

Toose, L., Woodfine, D.G., MacLeod, M., Mackay, D., Gouin, J. (2004). BETR-World: a geographically explicit model of chemical fate: application to transport of alpha-HCH to the Arctic Environmental Pollution 128, 223-40.

Van de Meent D. (1993). SimpleBox: A Generic Multi-media Fate Evaluation Model. National Institute for Public Health and the Environment. RIVM Report 672720 001. Bilthoven, NL.

Van de Meent, D., McKone, T.E., Parkerton, T., Matthies, M., Scheringer, M., Wania, F., Purdy, R., Bennett, D. (2000). Persistence and transport potential of chemicals in a multimedia environment. In: Klecka, g. et al. (Eds.) Evaluation of Persistence and Long-Range Transport Potential of Organic Chemicals in the Environment. SETAC Press, Pensacola FL, Chapter 5, pp. 169-204.

Van de Meent, D., Hollander, A., Peijnenburg, W., Breure, T. (2011). Fate and transport of contaminants. In: Sánches-Bayo, F., Van den Brink, P.J., Mann, R.M. (eds.),Ecological Impacts of Toxic Chemicals, Bentham Science Publishers, pp. 13-42.

Wania, F. (1999). On the origin of elevated levels of persistent chemicals in the environment. Environmental Science and Pollution Research 6, 11-19.

Wegmann, F., Scheringer, M., Hungerbühler, K. (2005). First investigations of mountainous cold condensation effects with the CliMoChem model. Ecotoxicology and Environmental Safety 63, 42-51.

3.8.2. Question 1

Describe, using your own words, the essential characteristics of mass balance equations.

3.8.2. Question 2

What is "steady state"? What is "equilibrium"? Use a few lines of text to describe the essentials, indicating differences and commonalities.

3.8.2. Question 3

Mass balance equations are used in models to calculate concentrations of substances in the environment, given knowledge of the rates of emission. Give a worked-out example for a one-compartment situation, e.g. a fresh-water lake.

3.8.2. Question 4

Name and describe one (or more) example(s) of multimedia mass balance modeling.

3.8.3. Metal speciation models

Authors: Wilko Verweij

Reviewers: John Parsons, Stephen Lofts

Learning objectives:

You should be able to

- Understand the basics of speciation modeling
- Understand the factors determining speciation and how to calculate them
- Understand in which types of situations speciation modeling can be helpful

Keywords: speciation modeling, solubility, organic complexation

Introduction

Speciation models allow users to calculate the speciation of a solution rather than to measure it in a chemical way or to assess it indirectly using bioassays (see section 3.5). As a rule, speciation models take total concentrations as input and calculate species concentrations.

Speciation models use thermodynamic data about chemical equilibria to calculate the speciation. This data, expressed in free energy or as equilibrium constants, can be found in the literature. The term 'constant' is slightly misleading as equilibrium constants depend on the temperature and ionic strength of the solution. The ionic strength is calculated from the concentrations (C) and charges (Z) of ions in solution using the equation:

$$1 = \frac{1}{2}\sum_{i=1}^{n} C_i Z_i^2$$

For many equilibria, no information is available to correct for temperature. To correct for ionic strength, many semi-empirical methods are available, none of which is perfect.

How these models work

For each equilibrium reaction, an equilibrium constant can be defined. For example, for the reaction

$Cu^{2+} + 4\ Cl^{-} \rightleftharpoons CuCl_4^{2-}$

the equilibrium constant can be defined as

$$\beta = \frac{[CuCl_4^{2-}]}{[Cu^{2+}] * [Cl^{-}]^4}$$

Consequently, when the concentrations of free Cu^{2+} and free Cl^{-} are known, the concentration of $CuCl_4^{2-}$ can be easily calculated as:



$[CuCl_4^{2-}] =$ $\beta * [Cu^{2+}] * [Cl^{-}]^4$

In fact, the concentrations of free Cu^{2+} and free Cl^{-} are often NOT known, but what is known are the total concentrations of Cu and Cl in the system. In order to find the speciation, we need to set up a set of mass balance equations needs to be set up, for example:

[total Cu] = [free Cu^{2+}] + [$CuOH^{+}$] + [$Cu(OH)_2$] + [$Cu(OH)_3^{-}$] (..) + [$CuCl^{+}$] + [$CuCl_2$] (..) etc.

[total Cl] = (..)

Each concentration of a complex is a function of the free concentrations of the ions that make it up. So we can say that if we know the concentrations of all the free ions, we can calculate the concentrations of all the complexes, and then we can calculate the total concentrations. A solution to the problem cannot be found by rearranging the mass balance equations, because they are non-linear. What a speciation model does is to repeatedly estimate the free ion concentrations, on each loop adjusting them so that the calculated total concentrations more closely match the known totals. When the calculated and known total concentrations all agree to within a defined precision, the speciation has been calculated. The critical part of the calculation is adjusting the free ion

concentrations in a sensible and efficient way to find the solution as quickly as possible. Several more or less sophisticated methods are available to solve this, but usually a Newton-Raphson method is applied.

Influence of temperature and ionic strength

In fact the explanation above is too simple. Equilbrium constants are valid under specific conditions for temperature and ionic strength (for example the standard conditions of 25^{o}C and



[endif]--> and need to be converted to the temperature and ionic strength of the system for which speciated is being calculated. It is possible to adapt the equilibrium constants for non-standard temperatures, but this requires knowledge of heat capacity (ΔH) data of each equilibrium. That knowledge is often not available. Constants can be converted from 25°C to other temperatures using the Van 't Hoff-equation:

$$ln(\frac{K_2}{K_1}) = -\frac{\Delta H}{R}(\frac{1}{T_2} - \frac{1}{T_1})$$

where K_1 and K_2 are the constants, T_1 and T_2 the temperatures, ΔH is the enthalpy of a reaction and R is the gas constant.

Equilbrium constants are also valid for one specific value of ionic strength. For conversion from one value of ionic strength to another, many different approaches may be used. This conversion is quite important, because already at relatively low ionic strengths, deviations from ideality become significant, and the activity of a species starts to deviate from its concentration. Hence, the intrinsic, or thermodynamic, equilibrium constants (i.e. constants at a hypothetical ionic strength of zero) are no longer valid and the activity a of ions at non-zero ionic strength needs to be calculated from the concentration and the activity coefficient:

$a = \gamma * c$

where γ is the activity coefficient (dimensionless; sometimes also called f) and c is the concentration; a and c are in mol/liter.

The first solution to calculate activity coefficients for non-zero ionic strength was proposed by Debye and Hückel in 1923. The Debye-Hückel theory assumes ions are point charges so it does not take into account the volume that these ions occupy nor the volume of the shell of ligands and/or water molecules around them. The Debye-Hückel gives good approximations, up to circa 0.01 M for a 1:1-electrolyte, but only up to circa 0.001 M for a 2:2-electrolyte. When the ionic strength exceeds these values, the activity coefficients that the Debye-Hückel approximation predicts deviate significantly from experimental values. Many environmental applications require conversions for higher ionic strengths making the Debye-Hückel-equation insufficient. To overcome this problem, many researchers have suggested other methods, like the extended Debye-Hückel-equation, the Güntelberg-equation and the Davies-equation, but also the Bromley-equation, the Pitzer-equation and the Specific Ion Interaction Theory (SIT).

Many programs use the Davies-equation, which calculates activity coefficients γ as follows:

$$log\ \gamma = -0.5079\ z^2\ (\frac{\sqrt{I}}{(1+\sqrt{I})} - 0.3\ I)$$

where z is the charge of the species and I the ionic strength. Sometimes 0.2 instead of 0.3 is used. Basically all these approaches take the Debye-Hückel-equation as a starting point, and add one or more terms to correct for deviations at higher ionic strengths. Although many of these methods are able to predict the activity of ions fairly well, they are in fact mainly *empirical* extensions without a solid theoretical basis.

Solubility

Most salts have a limited solubility; in several cases the solubility is also important under conditions that occur in the environment. For instance, for $CaCO_3$ the solubility product is $10^{-8.48}$, which means that when $[Ca^{2+}] * [CO_3^{2-}] > 10^{-8.48}$, $CaCO_3$ will precipitate, until $[Ca^{2+}] * [CO_3^{2-}] = 10^{-8.48}$. But it also works the other way around: if solid $CaCO_3$ is present in a solution where $[Ca^{2+}] * [CO_3^{2-}] < 10^{-8.48}$ (note the '<'-sign), solid $CaCO_3$ will dissolve, until $[Ca^{2+}] * [CO_3^{2-}] = 10^{-8.48}$. Note that the Ca and CO_3 in the formula here refer to *free* ions. For example, a 10^{-13} M solution of Ag_2S will lead to precipitation of Ag_2S. The *free* concentrations of Ag and S are $6.5*10^{-15}$ M and $1.8*10^{-22}$ M resp. (which corresponds with the solubility product of $10^{-50.12}$, but the *dissolved* concentrations of Ag and S are $7.1*10^{-15}$ M and $3.6*10^{-15}$ M resp., so for S seven order of magnitude higher. This is caused by the formation of S-complexes with protons (HS^- and H_2S *(aq)*) and to a lesser extent with Ag.

Complexation by organic matter

Complexation with Dissolved Organic Carbon (DOC) is different from inorganic complexation or complexation with well-defined compounds such as acetate or NTA. The reasons for that difference are as follows.

- DOC is very heterogeneous; DOC isolated at two sites may be very different (not to mention the difficulty of selecting isolation procedures).
- Complexation with DOC generally shows a continous range of equilibrium constants, due to chemical and steric differences in neighbouring groups.
- Increased cation binding and/or the ionic strength of the solution change electrostatic interactions among the functional groups in DOC-molecules, which influences the equilibrium constants.
- In addition, changing electrostatic interactions may cause conformational changes of the molecules.

Among the most popular models to assess organic complexation are Model V (1992), VI (1998) and VII (2011), also known as WHAM, written by Tipping and co-authors (Tipping & Hurley, 1992; Tipping, 1994, 1998; Tipping, Lofts & Sonke, 2011). All these models assume that two types of binding occur: *specific binding* and accumulation in the *diffuse double layer*. Specific binding is the formation of a chemical bond between an ion and a functional group (or groups) on the organic molecule. Diffuse double layer accumulation is the accumulation of ions of opposite electrical charge adjacent to the molecule, without formation of a chemical bond (the electrical charge is usually negative, so the ions that accumulate are cations).

For *specific binding*, all these models distinguish fulvic acids (FA) and humic acids (HA) which are treated separately. These two classes of DOC are typically the most abundant components of natural organic matter in the environment - in surface freshwaters, the fulvic acids are typically the most abundant. For each class, eight different discrete binding sites are used in the model. The sites have a range of acid-base properties. Metals bind to these sites, either to one site alone (monodentate), to two sites (bidentate) or, starting with Model VI, to three (tridentate). A fraction of the sites is allowed to form bidentate complexes. Starting with Model

VI, for each bidentate and tridentate group three sub-groups are assumed to be present - this further increases the range of metal binding strengths.

Binding constants depend on ionic strength and electrostatic interactions. Conditional constants are calculated in the same way in Model V, VI and VII, as follows:

$$K(z) = K * e^{2wZ}$$

where:

- Z is the charge of the organic acid (in moles per gram organic matter);
- w is calculated by :

$$w = P * log_{10}(I)$$

where:

- *P* is a constant term (different for FA and HA, and different for each model);
- *I* is the ionic strength.

Therefore, the conditional constant depends on the charge on the organic acids as well as on the ionic strength. For the binding of metals, the calculation of the conditional constant occurs in a similar way.

The *diffuse double layer* is usually negatively charged, so it is usually populated by cations, in order to maintain electric neutrality. Calculations for the diffuse double layer are the same for Model V, Model VI and Model VII. The volume of the diffuse double layer is calculated separately for each type of acid, as follows:

$$V_D = \frac{10N_{Av}}{M} * \frac{4\pi}{3} * [(r + \frac{1}{\kappa})^3 - r^3]$$

where:

- N_{Av} is Avogadro's number;
- *M* is the molecular weight of the acid;
- *r* is the radius of the molecule (0.8 nm for fulvic acids, 1.72 for humic acids);
- κ is the Debye-Hückel parameter, which is dependent on ionic strength.

Simply applying this formula in situations of low ionic strength and high content of organic acid would lead to artifacts (where volume of diffuse layer can be calculated to be more than 1 liter/liter). Therefore, some "tricks" are implemented to limit the volume of the diffuse double layer to 25% of the total.

In case the acid has a negative charge (as it has in most cases), positive and neutral species are allowed to enter the diffuse double layer, just enough to make the diffuse double layer electrically neutral. When the acid has a positive charge, negative and neutral species are present.

The concentration of species in the diffuse double layer is calculated by assuming that the concentration of that species in the diffuse double layer depends on the concentration in the bulk solution and the charge.

In formula:

$$\frac{[X^Z]_{DDL}}{[X^Z]_{solution}} = R^{|Z|}$$

where R is calculated iteratively, to ensure the diffuse double layer is electrically neutral.

Applications

Speciation models can be used for many purposes. Basically, two groups of applications can be distinguished. The first group consists of applications meant to understand the chemical behaviour of any system. The second group focuses on bioavailability.

Chemical behaviour; laboratory situations

Speciation models can be helpful in understanding chemical behaviour in either laboratory situations or field situations. For instance, if you want to add EDTA to a solution to prevent metals from precipitation, the choice of the EDTA-substance also

determines the pH of the final solution. Figure 1 shows the pH of a 1 mM solution of EDTA for five different EDTA-salts. This shows that if you want to end up with a near neutral solution, the best choice is to add EDTA as the Na_3HEDTA-salt. Adding a different salt requires adding either acid or base, or more buffer capacity, which in turn will influence the chemical behaviour of the solution.

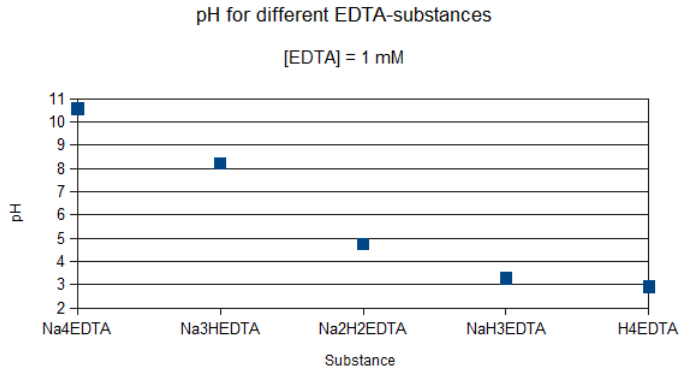


Figure 1. pH of a 1 mM EDTA-solution for different EDTA-salts. Data obtained using speciation program CHEAQS Next.

If you have field measurements of redox potential, speciation models can help to predict whether iron will be present as Fe(II) or Fe(III), which is important because Fe(II) behaves quite different chemically than Fe(III) and also has a quite different bioavailability. The same holds for other elements that undergo redox equilibria like N, S, Cu or Mn.

Phase reactions can be predicted with speciation models, for example the dissolution of carbonate due to the gas solution reaction of CO_2. Another example is the speciation in Dutch Standard Water (DSW), a frequently used test medium for ecotoxicological experiments, which is oversaturated with respect to $CaCO_3$ and therefore displays a part of Ca as a precipitate. The fraction that precipitates is very small (less than 2% of the Ca) so it seems unimportant at first glance, but the precipitate induces a pH-shift of 0.22, a factor of almost two in the concentration of free H^+.

Many metals are amphoteric and have therefore a minimum solubility at a moderate pH, while dissolving more at both higher and lower pH-values. This can easily be seen in the case of Al: Figure 2 shows the concentration of dissolved Al as a function of pH (note log-scale for Y-axis). Around pH of 6.2, the solubility is at its minimum. At higher and lower pH-values, the solubility is (much) higher.

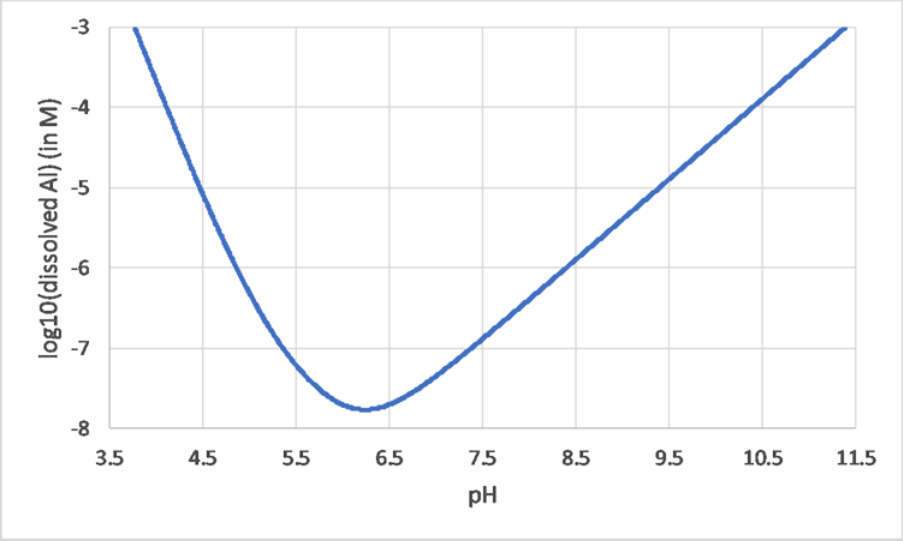


Figure 2. Soluble Al as function of pH. Data obtained using speciation program CHEAQS Next.

Speciation models can also help to understand differences in the growth of organisms or adverse effects on organisms, in different chemical solutions. For example, Figure 3 shows that changes in speciation of boron can be expected only between roughly pH 8 and 10.5, so when you observe a biological difference between pH 7 and 8, it is not likely that boron is the cause. Copper on the

other hand (see Figure 4) does display differences in speciation between pH 7 and 8 so is a more likely cause of different biological behaviour.

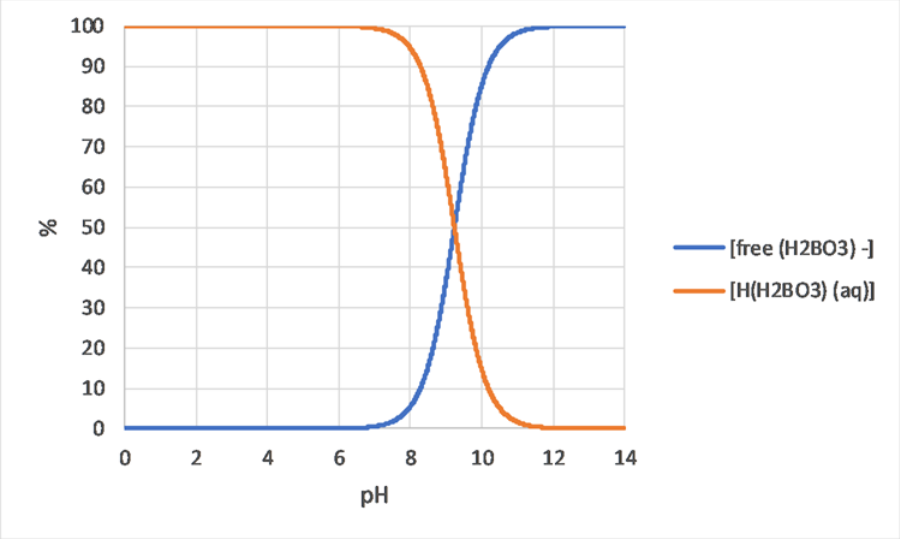


Figure 3. Speciation of B as function of pH; concentration of boron was $1x10^{-6}$ M. At higher concentrations, complexes with 2, 3, 4 or 5 B-ions can be formed at significant concentrations. Data obtained using speciation program CHEAQS Next.

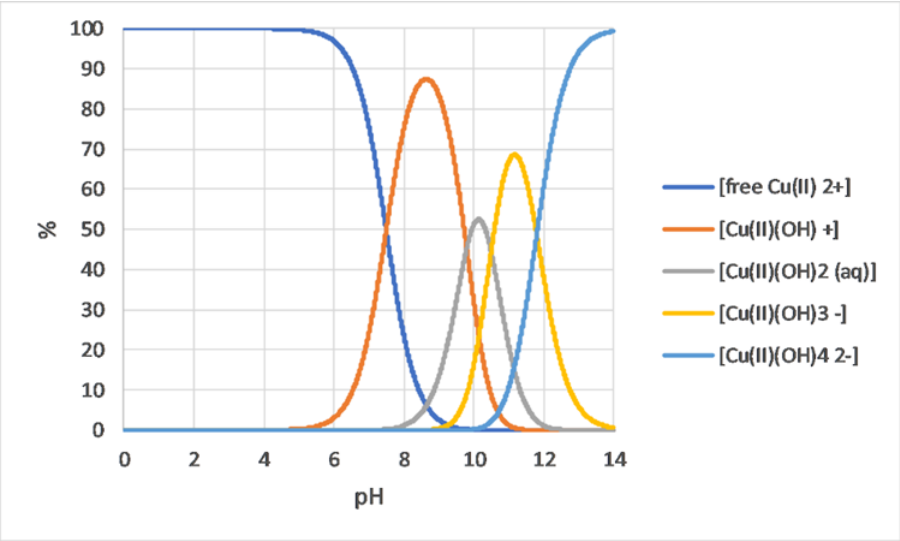


Figure 4. Speciation of copper(II) as a function of pH; concentration of copper was $3x10^{-8}$ M. At higher concentrations, complexes with 2 or 3 Cu(II)-ions can be formed at significant concentrations. Data obtained using speciation program CHEAQS Next.

Chemical behaviour: field situations

In field situations, the chemistry is usually much more complex than under laboratory conditions. Decomposition of organisms (including plants) results in a huge variety of organic compounds like fulvic acids, humic acids, proteins, amino acids, carbohydrates, etc. Many of these compounds interact strongly with cations, some also with anions or uncharged molecules. In addition, metals easily adsorb to clay and sand particles that are found everywhere in nature. To make it more complex, suspended matter can contain a high content of organic material which is also capable of binding cations.

For complexation by fulvic and humic acids, Tipping and co-workers have developed a unifying model (Tipping & Hurley, 1992; Tipping, 1994, 1998; Tipping, Lofts & Sonke, 2011). The most recent version, WHAM 7 (Tipping, Lofts & Sonke, 2011), is able to predict cation complexation by fulvic acids and humic acids over a wide range of chemical circumstances, despite the large difference in composition of these acids. This model is now incorporated in several speciation programs.

Suspended matter may be of organic or of inorganic character. *Inorganic matter* usually consists of (hydr)oxides of metals, such as Mn, Fe, Al, Si or Ti, and clay minerals. In practice, the (hydr)oxides and clays occur together, but the mutual proportions may differ dramatically depending on the source. Since the chemical properties of these metal (hydr)oxides clays are quite different, there is a huge variation in the chemical properties of inorganic suspended matter in different places and different times. As a consequence, modeling interactions between dissolved constituents and suspended inorganic matter is challenging. Only by measuring some properties of suspended inorganic matter, can modeling be applied successfully. For suspended *organic matter*, the variation in properties is also large and modelling is challenging.

Bioavailability

Speciation models are useful in understanding and assessing the bioavailability of metals and other elements in test media. Test media often contain substances like EDTA to keep metals in solution. EDTA-complexes in general are not bioavailable, so in addition to keeping metals in solution they also change their bioavailability. Models can calculate the speciation and help you to assess what is actually happening in a test medium. An often forgotten aspect is the influence of CO_2. CO_2 from the ambient atmosphere can enter a solution or carbonate in solution (if in excess over the equilibrium concentration) can escape to the atmosphere. The degree to which this exchange takes place, influences the pH of the solution as well as the amount of carbonate that stays in solution (carbonates are often poorly soluble).

Similarly, in field situations models can help to understand the bioavailability of elements. As stated above, the influence of DOC can nowadays be assessed properly in many situations, the influence of suspended matter remains more difficult to assess. Nevertheless models can deliver insights in seconds that otherwise can be obtained only with great difficulty.

Models

There are many speciation programs available and several of them are freely available. Usually they take a set of total concentrations as input, plus information about parameters such as pH, redox, concentration of organic carbon etc. Then the programs calculate the speciation and present them to the user. The equations cannot be solved analytically, so an iterative procedure is required. Although different numerical approaches are used, most programs construct a set of non-linear mass balance equations and solve them by simple or advanced mathematics. A complication in this procedure is that the equilibrium constants depend on the ionic strength of the solution, and that this ionic strength can only be calculated when the speciation is known. The same holds for the precipitation of solids. The procedure is shown in Figure 5.

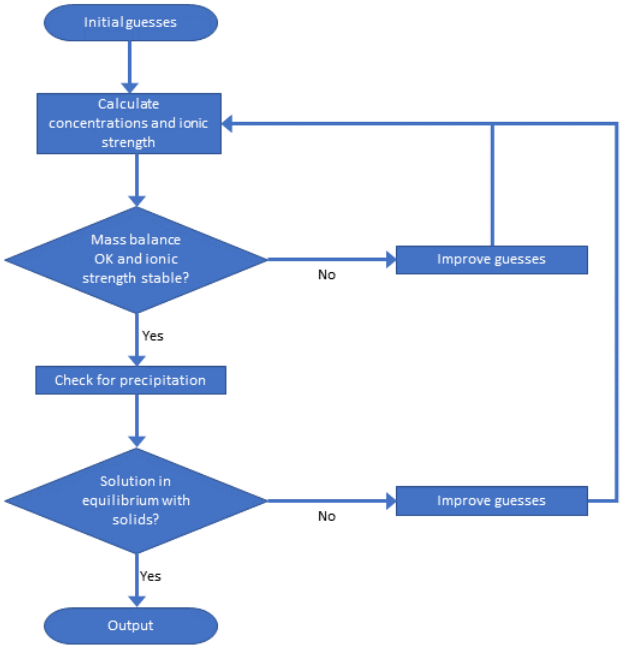


Figure 5. Typical flow diagram of a speciation program.

Limitations

For modeling speciation, thermodynamic data is needed for all relevant equilibrium reactions. For many equilibria, this information is available, but *not for all*. This hampers the usefulness of speciation modeling. In addition, there can be *large variations* in the thermodynamic values found in the literature, resulting in uncertainty about the correct value. A factor of 10 between the highest and lowest values found is not an exception. This of course influences the reliability of speciation calculations. For many equilibria, the thermodynamic data is only available for the *standard temperature* of 25°C and no information is available to assess the data at other temperatures, although the effect of temperature can be quite strong. Also ionic strength has a high impact on equilibrium 'constants'; there are many methods available to correct for the effect of ionic strength, but most of them are at best semi-empirical. Simonin (2017) recently proposed a method with a solid theoretical basis; however, the data required for his method are available only for a few complexes so far.

More fundamentally, you should realize that speciation programs typically calculate the *equilibrium situation*, while some reactions are very slow and, more inportant, nature is in fact a very dynamic system and therefore never in equilibrium. If a system is close to equilibrium, speciation programs can often make a good assessment of the actual situation, but the more dynamic a system, the more care you should take in believing the programs' results. Nevertheless it is good to realise that a chemical system will *always* move towards the equilibrium situation, while organisms may move them away from equilibrium. Phototrophs are able to move a system *away from* its equilibrium situation whereas decomposers and heterotrophs generally help to move a system towards its equilibrium state.

References

Simonin , J.-P. (2017). Thermodynamic consistency in the modeling of speciation in self-complexing electrolytes. Ind. Eng. Chem. Res. 56, 9721-9733.

Tipping, E., Hurley, M.A. (1992). A unifying model of cation binding by humic substances. Geochimica et Cosmochimica Acta 56, 3627 - 3641.

Tipping, E. (1994). WHAM - A chemical equilibrium model and computer code for waters, sediments, and soils incorporating a discrete site/electrostatic model of ion-binding by humic substances. Computers & Geosciences 20, 973 - 1023.

Tipping, E. (1998). Humic Ion-Binding Model VI: An Improved Description of the Interactions of Protons and Metal Ions with Humic Substances. Aquatic Geochemistry 4, 3 - 48.

Tipping, E., Lofts, S., Sonke, J.E. (2011). Humic Ion-Binding Model VII: a revised parameterisation of cation-binding by humic substances. Environmental Chemistry 8, 228 - 235.

Further reading

Stumm, W., Morgan, J.J. (1981). Aquatic chemistry. John Wiley & Sons, New York.

More, F.M.M., Hering, J.G. (1993). Principles and Applications of Aquatic Chemistry. John Wiley & Sons, New York.

3.8.3. Question 1

Briefly describe what speciation models are.

3.8.3. Question 2

What are the factors determining speciation and how can these be accounted for?

3.8.3. Question 3

Give two examples of situations where speciation modeling can be useful.

3.8.4. Modeling exposure at ecological scales

In preparation

This page titled 2.3.4.7: Modelling exposure is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.7: Modelling exposure** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.4.8: Section 4.1. Toxicokinetics

Section 4.1. Toxicokinetics

If you want to re-use this section, for e.g. in your electronic learning environment, feel free to copy this url: https://maken.wikiwijs.nl/162476/4_1__Toxicokinetics

---

This page titled 2.3.4.8: Section 4.1. Toxicokinetics is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.8: Section 4.1. Toxicokinetics** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.4.9: Section 4.4. Increasing ecological realism in toxicity testing

Section 4.4. Increasing ecological realism in toxicity testing

If you want to re-use this section, for e.g. in your electronic learning environment, feel free to copy this url: https://maken.wikiwijs.nl/162480/4_4__Increasing_ecological_realism_in_toxicity_testing

---

This page titled 2.3.4.9: Section 4.4. Increasing ecological realism in toxicity testing is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.9: Section 4.4. Increasing ecological realism in toxicity testing** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.4.10: Section 4.2. Toxicodynamics and Molecular Interactions

Section 4.2. Toxicodynamics & Molecular Interactions

If you want to re-use this section, for e.g. in your electronic learning environment, feel free to copy this url: https://maken.wikiwijs.nl/162477/4_2__Toxicodynamics___Molecular_Interactions

---

This page titled 2.3.4.10: Section 4.2. Toxicodynamics and Molecular Interactions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.10: Section 4.2. Toxicodynamics and Molecular Interactions** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.4.11: Section 4.3. Toxicity testing

Section 4.3. Toxicity testing

This page titled 2.3.4.11: Section 4.3. Toxicity testing is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **4.11: Section 4.3. Toxicity testing** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

SECTION OVERVIEW

2.3.5: Population, Community and Ecosystem Ecotoxicology

2.3.5.1: Introduction- Linking population, community and ecosystem responses

2.3.5.2: Population ecotoxicology in laboratory settings

2.3.5.3: Wildlife population ecotoxicology

2.3.5.4: Trait-based approaches

2.3.5.5: Population models

2.3.5.6: Metapopulations

2.3.5.7: Community ecotoxicology

2.3.5.8: Structure versus function incl. ecosystem services

2.3.5.9: Landscape ecotoxicology

This page titled 2.3.5: Population, Community and Ecosystem Ecotoxicology is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

2.3.5.1: Introduction- Linking population, community and ecosystem responses

5.1. Introduction: Linking population, community and ecosystem responses

In preparation

This page titled 2.3.5.1: Introduction- Linking population, community and ecosystem responses is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **5.1: Introduction- Linking population, community and ecosystem responses** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.5.2: Population ecotoxicology in laboratory settings

5.2. Population ecotoxicology in laboratory settings

Author: Michiel Kraak

Reviewers: Nico van den Brink and Matthias Liess

Learning objectives:

You should be able to

· motivate the importance of studying ecotoxicology at the population level.

· name the properties of populations, unique to this level of biological organisation.

· explain the implications of age and developmental stage specific sensitivities for population responses to toxicant exposure.

Key words: Population ecotoxicology, density, age structure, population growth rate

Introduction

The motivation to study ecotoxicological effects at the population level is that generally the targets of environmental protection are indeed populations, communities and ecosystems. Additionally, several phenomena are unique to this level, including age specific sensitivity and interaction between individuals. Studying the population level is distinguished from the individual level and lower by a less direct link between the chemical exposure and the observed effects, due to individual variability and several feedback loops, loosening the dose-response relationships. Research at the population level is thus characterized by an increasing level of uncertainty if these processes are not properly addressed and by increasing time and efforts. Hence, it is not surprising that effects at the population are understudied. This is even more the case for investigations on higher levels like meta-populations, communities and ecosystems (see sections on meta-populations, communities and ecosystems). It is thus highly important to obtain data and insights into mechanisms leading to effects at the population level, keeping in mind the relevant interactions with lower and higher levels of organisation.

Properties of populations are unique to this level of biological organization and include social structure (see section on invertebrate community ecotoxicology), genetic composition (see section on genetic variation), density and age structure. This gives room to age and developmental stage specific sensitivities to chemicals. For almost all species, young individuals like neonates or first instars are markedly more sensitive than adults or late instar larvae. This difference may run up to three orders of magnitude and consequently instar specific sensitivities may vary as much as species specific sensitivities (Figure 1). Population developmental stage specific sensitivities have also been reported. Exponentially growing daphnid populations exposed to the insecticide fenvalerate recovered much faster than populations that reached carrying capacity (Pieters and Liess, 2006). Given the age and developmental stage specific sensitivities, the timing of exposure to toxicants in relation to the critical life stage of the organism may seriously affect the extent of the adverse effects, especially in seasonally synchronised populations.

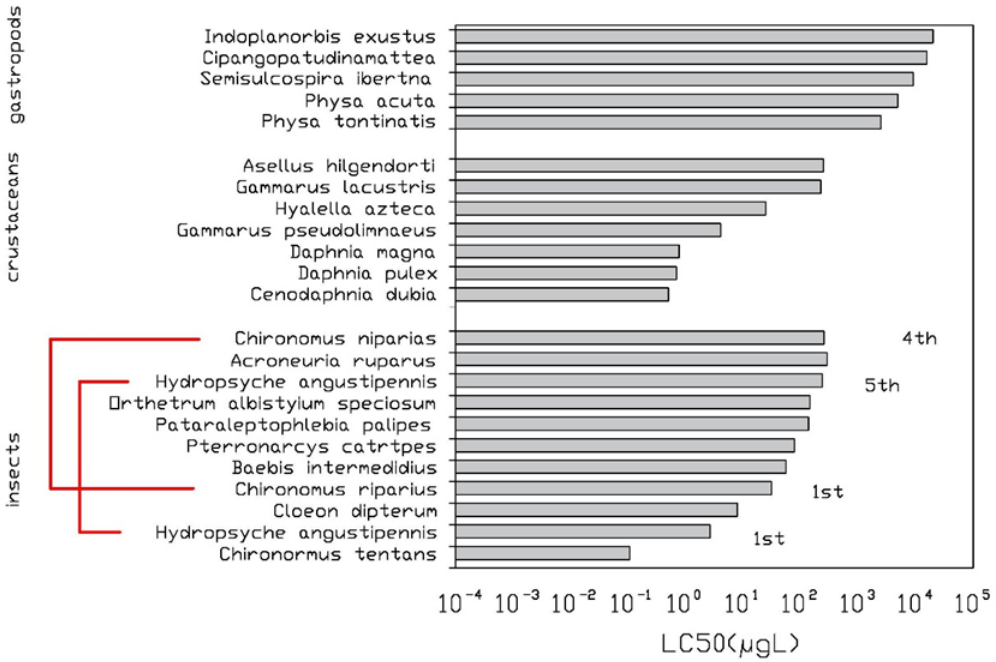


Figure 1. 48h LC50 values of the insecticide diazinon for insects, crustaceans, and gastropods ranked according to sensitivity (according to Stuijfzand et al., 2000), showing that instar specific sensitivities may vary as much as species specific sensitivities. Drawn by Wilma IJzerman.
Note: this figure is a draft that still needs correction

A challenging question involved in population ecotoxicology is when a population is considered to be stable or in steady state. In spite of the various types of oscillation all populations depicted in Figure 2 can be considered to be stable. One could even argue that any population that does not go extinct can be considered stable. Hence, a single population could vary considerable in density over time, potentially strongly affecting the impact of exposure to toxicants.

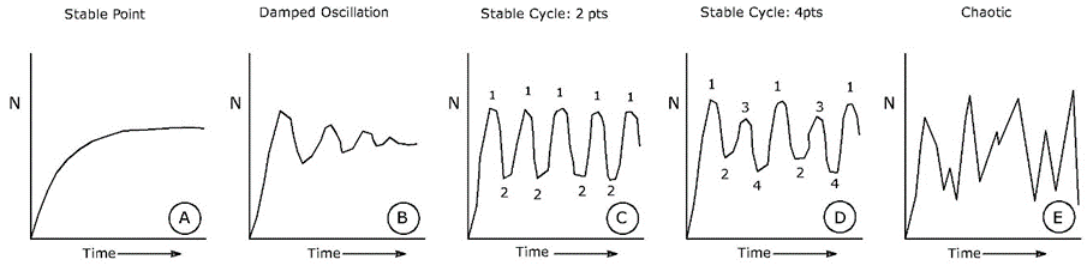


Figure 2. Different types of population development over time. Drawn by Wilma IJzerman.

When populations suffer from starvation and crowding due to high densities and intraspecific competition, they are markedly more sensitive to toxicants, sometimes even up to a factor of 100 (Liess et al., 2016). This may even lead to unforeseen, indirect effects. Relative population growth rate (individual/individual/day) of high density populations of chironomids actually increased upon exposure to Cd, because Cd induced mortality diminished the food shortage for the surviving larvae (Figure 3). Only at the highest Cd exposure population growth rate decreased again. For populations at low densities, the anticipated decrease in population growth rate with increasing Cd concentrations was observed. Yet, at all Cd exposure levels growth rate of low density populations was markedly higher than that of high density populations.

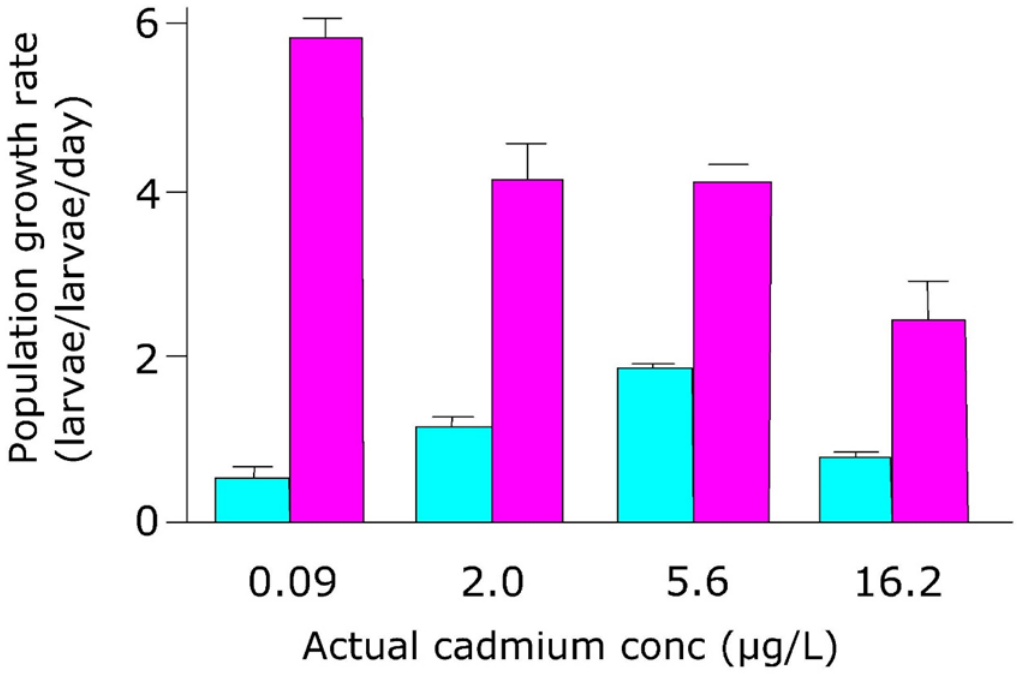


Figure 3. Effects of cadmium exposure and density on population growth rate of Chironomus riparius (according to Postma et al., 1994). Mean values with standard error. Blue bars represent the high larval density and purple bars the low larval density. Redrawn by Wilma IJzerman.

Population ecotoxicity tests

In *chronic* ecotoxicity studies, preferably cohorts of individuals of the same size and age are selected to minimize variation in the outcome of the test, whereas in *population* ecotoxicology the natural heterogenous population composition is taken into account. This does make it harder though to interpret the obtained experimental data. Especially when studying populations of higher organisms in the wild, the increasing time to complete the research due to the long life span of these organisms imposes practical limitations (see section on wildlife population ecotoxicology). In the laboratory, this can be circumvented by selecting test species with relatively short life cycles, like algae, bacteria and zooplankton. For algae, a three or four day test can be considered as a multigeneration experiment and during 21 d female daphnids may release up to three clutches of neonates. These population ecotoxicity tests offer the unique possibility to calculate the ultimate population parameter, the population growth rate (r). This is a demographic population parameter, integrating survival, maturity time and reproduction (see section on population modeling). Yet, such chronic experiments are typically performed with cohorts and not with natural populations, making these experiments rather an extension of chronic toxicity tests than true population ecotoxicity tests.

References

Knillmann, S., Stampfli, N.C., Beketov, M.A., Liess, M. (2012). Intraspecific competition increases toxicant effects in outdoor microcosms. Ecotoxicology 21, 1857-1866.

Liess, M., Foit, K., Knillmann, S., Schäfer, R.B., Liess, H.-D. (2016). Predicting the synergy of multiple stress effects. Scientific Reports 6, 32965.

Pieters, B.J., Liess, M. (2006). Population developmental stage determines the recovery potential of *Daphnia magna* populations after fenvalerate application. Environmental Science and Technology 40, 6157-6162.

5.2. Question 1

Motivate the importance of studying ecotoxicology at the population level and higher.

5.2. Question 2

Name the properties of populations that are unique to this level of biological organisation.

5.2. Question 3

Why is it important to understand the implications of age and developmental stage specific sensitivities for population responses to toxicant exposure?

5.2. Question 4

Explain the results observed in Figure 3.

This page titled 2.3.5.2: Population ecotoxicology in laboratory settings is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **5.2: Population ecotoxicology in laboratory settings** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.5.3: Wildlife population ecotoxicology

5.3. Wildlife population ecotoxicology

5.3.1. Forensic investigation into crash of Asian vulture populations

Author: Nico van den Brink

Reviewers: Ansje Löhr, John Elliott

Learning objectives:

You should be able to

- describe how forensic approaches are used in ecotoxicology
- critically reflect on the uncertainty of prospective risk assessment of new chemicals

Keywords: Pharmaceuticals, uncertainty, population decline, retrospective monitoring

Introduction

Historically, vulture populations in India, Pakistan and Nepal were too numerous to be effectively counted. In the mid-1990s numbers in northern India started to decline catastrophically, which was evidenced in the Keoladeo National Park (figure 1, Prakash 1999). Further monitoring of population numbers indicated unprecedented declines of over 90-99% from the mid-1990s to the early 2000s for Oriental White-backed vultures (*Gyps bengalensis*), Long-billed vultures (*Gyps indicus*) and also Slender-billed vultures (*Gyps tenuirostris*) (Prakash 1999).

Figure 1. Populations of White-backed vultures in Keoladeo National park in different years. Redrawn from Prakash (1999) by Wilma IJzerman.

In the following years, similar declines were observed in Pakistan and Nepal, indicating that the causative factor was not restricted to a specific country or area. Total losses of vultures were estimated to be in the order of tens of millions. The first ideas about potential causes of those declines focussed on known infectious diseases or the possibility of new diseases to which the vulture population had not been previously exposed. However, no diseases were identified that had shown similar rates of mortalities in other bird species. Vultures are also considered to have a highly developed immune response given their diet of scavenging dead and often decaying animals. To obtain insights, initial interdisciplinary ecological studies were performed to provide a basic understanding of background mortality in the species affected. These studies started in large colonies in Pakistan, but were literally races against time, as some populations had already decreased by 50%, while others were already extirpated, (Gilbert et al., 2006). Despite those difficulties it was determined that mortalities were occurring principally in adult birds and not at the nestling phase. More in depth studies were performed to discriminate between natural mortality of for instance juvenile fledglings, which may be high in summer, just after fledging. After scrutinising the data no seasonality was observed in the abnormal, high mortality, indicating that this was not related to breeding activities. The investigations also revealed another important factor that these vultures were predominantly feeding on domestic livestock, while telemetric observations, using transmitters to assess flight and activity patterns f the birds, showed that the individual birds could range over very long distances to reach carcasses of livestock (up to over 100 km).

Since no apparent causes for mortality were obtained in the ecological studies, more diagnostic investigations were started, focussing on infectious diseases and carried out in Pakistan (Oaks, 2011). However, that was easier said than done. Since large numbers of birds died, it was deemed essential to establish the logistics necessary to perform the diagnostics, including post-mortems, on all birds found dead. Although high numbers of birds died, hardly any fresh carcasses were available, due to remoteness of some areas, the presence of other scavengers and often hot conditions which fostered rapid decay of carcasses. Post-mortems on a selection off birds revealed that birds suspect of abnormal mortality all suffered from visceral gout, which is a white pasty smear covering tissues in the body including liver and heart. In birds, this is indicative for kidney failure. Birds metabolise nitrogen into uric acid (mammals into urea) which is normally excreted with the faeces. However, in case of kidney failure the uric acid is not excreted but deposited in the body. Further inspections of more birds confirmed this, and the working hypothesis became that the increased mortality was caused by a factor inducing kidney failure in the birds.

Based on the establishment of kidney failure as the causative factor, histological and pathological studies were performed on several birds found dead which revealed that in birds with visceral gout, kidney lesions were severe with acute renal tubular necrosis (Oaks et al., 2004), confirming the kidney failure hypothesis. However, no indications of inflammatory cell infiltrations were apparent, ruling out the possibilities of infectious diseases. Those observation shifted the focus to potential toxic effects, although no previous case was known with a chemical causing such severe and extremely acute effects. First the usual suspects for kidney failure were addressed, like trace metals (cadmium, lead) but also other acute toxic chemicals like organophosphorus and carbamate pesticides and organochlorine chemicals None of those chemicals occurred at levels of concern and were ruled out. That left the researchers without leads to any clear causative factor, even after years of study!

Some essential pieces of information were available, however:

1) acute renal failure seemed associated with the mortality,

2) no infectious agent was likely to be causative pointing to chemical toxicity,

3) since exposure was likely to be via the diet the chemical exposure needed to be related to livestock (the predominant diet for the vultures), pointing to compounds present in livestock such as veterinarian products,

4) widespread use of veterinarian chemicals had started relatively recently.

After a survey of veterinarians in the affected areas of Pakistan, a single veterinarian pharmaceutical matched the criteria, diclofenac. This is a non-steroid anti-inflammatory drug (NSAID) since long used in human medicine but only introduced since the 1990s as a veterinarian pharmaceutical in India, Pakistan and surrounding countries. NSAIDs are known nephrotoxic compounds, although no cases were known with such acute and sever impacts. Chemical analyses of kidneys of vultures confirmed that kidneys of birds with visceral gout contained diclofenac, birds without signs of visceral gout did not. Also kidneys from birds that showed visceral gout and that died in captivity while being studied, were positive for diclofenac, as was the meat they were fed with. This all indicated diclofenac toxicity as the cause of the mortality, which was validated in exposure studies, dosing captive vultures with diclofenac. The species of *Gyps* vultures appeared extremely sensitive to diclofenac, showing toxic effects at 1% of the therapeutic dose for livestock mammalian species.

This underlying mechanism for that sensitivity has yet to be explained, but initially it was also unclear why the populations were impacted to such severe extent. That was found to be related to the feeding ecology of the vultures. They were shown to fly long ranges to search for carcasses, and as a result of that they show very aggregated feeding, i.e. a lot of birds on a single carcass (Green et al., 2004). Hence, a single contaminated carcass may expose an unexpected large part of the population to diclofenac. Hence, a combination of extreme sensitivity, foraging ecology and human chemical use caused the onset of extreme population declines of some Asian vulture species of the Gyps genus, or so called "Old World vultures".

This case demonstrated the challenges involved in attempting to disentangle the stressors causing very apparent population effects even on imperative species like vultures. It took several years of different groups of excellent researcher to perform the necessary research and forensic studies (under sometimes difficult conditions). Lessons learned are that even for compounds that have been used for a long time and thought to be well understood, unexpected effects may become evident. There is consensus that such effects may not be covered in current risk assessments of chemicals prior to their use and application, but this also draws attention to the need for continued post-market monitoring of organisms for potential exposure and effects. It should be noted that even nowadays, although the use of diclofenac is prohibited in larger parts of Asia, continued use still occurs due to its effectiveness in treating livestock and its low costs making it available to the farmers. Nevertheless, populations of Gyps vultures have shown to recover slowly.

References

Green, R.E., Newton, I.A., Shultz, S., Cunningham, A.A., Gilbert, M., Pain, D.J., Prakash, V. (2004). Diclofenac poisoning as a cause of vulture population declines across the Indian subcontinent. Journal of Applied Ecology 41, 793-800.

Gilbert, M., Watson, R.T., Virani, M.Z., Oaks, J.L., Ahmed, S., Chaudhry, M.J.I., Arshad, M., Mahmood, S., Ali, A., Khan, A.A. (2006). Rapid population declines and mortality clusters in three Oriental whitebacked vulture *Gyps bengalensis* colonies in Pakistan due to diclofenac poisoning. Oryx 40, 388-399.

Oaks, J.L., Gilbert, M., Virani, M.Z., Watson, R.T., Meteyer, C.U., Rideout, B.A., Shivaprasad, H.L., Ahmed, S., Chaudhry, M.J.I., Arshad, M., Mahmood, S., Ali, A., Khan, A.A. (2004). Diclofenac residues as the cause of vulture population decline in Pakistan. Nature 427, 630-633.

Oaks, J.L., Watson, R.T. (2011). South Asian vultures in crisis: Environmental contamination with a pharmaceutical. In: Elliott, J.E., Bishop, C.A., Morrissey, C.A. (Eds.) Wildlife Ecotoxicology. Springer, New York, NY. pp. 413-441.

Prakash, V. (1999). Status of vultures in Keoladeo National Park, Bharatpur, Rajasthan, with special reference to population crash in *Gyps* species. Journal of the Bombay Natural History Society 96, 365-378.

5.3.1. Question 1

Which ecological traits make vulture populations extremely vulnerable to exposure to chemicals like diclofenac?

5.3.1. Question 2

What indirect effects do you expect from the chemical induced crashes of the Asian vultures?

5.3.2. Otters, to PCB or not to PCB?

Author: Nico van den Brink

Reviewers: Ansje Löhr, Michiel Kraak, Pim Leonards, John Elliott

Learning objectives

You should be able to:

- explain the derivation of toxic threshold levels by extrapolating between species
- critically analyse implications of risk assessment for the conservation of species

Keywords: Threshold levels, read across, species specific sensitivity

The European otter (*Lutra lutra*) is a lively species which historically ranges all over Europe. In the second half of last century populations declined in North-West Europe, and at the end of the 1980s the species was declared extinct in the Netherlands. Several factors contributed to these declines, exposure to polychlorinated biphenyls (PCBs) and other contaminants was considered a prominent cause. PCBs can have different effects on organisms, primarily Ah-receptor mediated (see section on Receptor interactions). In order to assess the actual contribution of chemical exposure to the extinction of the otters, and the potential for population recovery it is essential to gain insight in the ratios between exposure levels and risk thresholds. However, since otters are rare and endangered, limited toxicological data is available on such thresholds. Most toxicological data is therefore inferred from research on another mustelids species the mink (*Mustela vison*) (Basu et al., 2007) a high trophic level, piscivorous species often used in toxicological studies. Several studies show that mink is quite sensitive to PCBs, showing e.g. effects on the length of the baculum of juveniles (Harding et al., 1999) and induction of hepatic enzyme systems and jaw lesions (Folland et al., 2016). Based on such studies, several threshold levels for otters were derived, depending on the toxic endpoints addressed. Based on number of offspring size and kit survival, EC_{50} were derived of approximately 1.2 to 2.4 mg/kg *wet* weight (Leonards et al., 1995), while for decreases in vitamin A levels due to PCB exposure, a safety threshold of 4 mg/kg in blood was assessed (Murk et al., 1998).

To re-establish a viable population of otters in the Netherlands, a program was established in the mid-1990s to re-introduce otters in the Netherlands, including monitoring of PCBs and other organic contaminants in the otters. Otters were captured in e.g. Belarus, Sweden and Czech Republic. Initial results showed that these otters already contained < 1 mg/kg PCBs based on wet weight (van den Brink & Jansman, 2006), which was considered to be below the threshold limits mentioned before. Individual otters were radio-tagged, and most were recovered later as victims of car incidences. Over time, PCB concentrations had changed, although not in the same direction for all specimen. Females with high initial concentrations showed declining concentrations, due to lactation, while in male specimens most concentrations increased over time, as you would expect. Nevertheless, concentrations were in the range of the threshold levels, hence risks on effects could not be excluded. Since the re-introduction program was established in a relatively low contaminated area in the Netherlands, questions were raised for re-introduction plans in more contaminated areas, like the Biesbosch where contaminants may still affect otters .

To assess potential risks of PCB contamination in e.g. the Biesbosch for otter populations a modelling study was performed in which concentrations in fish from the Biesbosch were modelled into concentrations in otters. Concentrations of PCBs in the fish differed between species (lipid rich fish such as eel greater concentrations than lean white fish), size of the fish (larger fish greater concentrations than smaller fish) and between locations within the Biesbosch. Using Biomagnification Factors (BMFs) specific for each PCB-congener (see section on Complex mixtures), total PCB concentrations in lipids of otters were calculated based on fish concentrations and different compositions of fish diets of the otters (e.g. white fish versus eel, larger fish versus smaller fish, different locations). Different diets resulted in different modelled PCB concentrations in the otters, however all modelled concentrations were above the earlier mentioned threshold levels (van den Brink and Sluiter, 2015). This would indicate that risks of effects for otters could not be ruled out, and led to the notion that release of otters in the Biesbosch would not be the best option.

However, a major issue related to such risk assessment is whether the threshold levels derived from mink are applicable to otter. The resulting threshold levels for otter are rather low and exceedance of these concentrations has been noticed in several studies. For instance, in well-thriving Scottish otter populations PCBs levels have been recorded greater than 50 mg/kg lipid weight in livers (Kruuk & Conroy, 1996). This is an order of magnitude higher than the threshold levels, which would indicate that even at higher concentrations, at which effects are to be expected based on mink studies, populations of free ranging otters do not seem to be affected adversely. Based on this, the applicability of mink-derived threshold levels for otters may be open to discussion.

The case on otters showed that the derivation of ecological relevant toxicological threshold levels may be difficult due to the fact that otters are not regularly used in toxicity tests. Application of data from a related species, in this case the American mink, however, may be limited due to differences in sensitivity. In this case, this could result in too conservative assessments of the risks, although it should be noted that this may be different in other combinations of species. Therefore, the read across of information of closely related species should therefore be performed with great care.

References

Basu, N., Scheuhammer, A.M., Bursian, S.J., Elliott, J., Rouvinen-Watt, K., Chan, H.M. (2007). Mink as a sentinel species in environmental health. Environmental Research 103, 130-144.

Harding, L.E., Harris, M.L., Stephen, C.R., Elliott, J.E. (1999). Reproductive and morphological condition of wild mink (*Mustela vison*) and river otters (*Lutra canadensis*) in relation to chlorinated hydrocarbon contamination. Environmental Health Perspectives 107, 141-147.

Folland, W.R., Newsted, J.L., Fitzgerald, S.D., Fuchsman, P.C., Bradley, P.W., Kern, J., Kannan, K., Zwiernik, M.J. (2016). Enzyme induction and histopathology elucidate aryl hydrocarbon receptor-mediated versus non-aryl receptor-mediated effects of Aroclor 1268 in American Mink (*Neovison vison*). Environmental Toxicology and Chemistry 35, 619-634.

Kruuk, H., Conroy, J.W.H. (1996). Concentrations of some organochlorines in otters (*Lutra lutra* L) in Scotland: Implications for populations. Environmental Pollution 92, 165-171.

Leonards, P.E.G., De Vries, T.H., Minnaard, W., Stuijfzand, S., Voogt, P.D., Cofino, W.P., Van Straalen, N.M., Van Hattum, B. (1995). Assessment of experimental data on PCB-induced reproduction inhibition in mink, based on an isomer- and congener-specific approach using 2,3,7,8-tetrachlorodibenzo-p-dioxin toxic equivalency. Environmental Toxicology and Chemistry 14, 639-652.

Murk, A.J., Leonards, P.E.G., Van Hattum, B., Luit, R., Van der Weiden, M.E.J., Smit, M. (1998). Application of biomarkers for exposure and effect of polyhalogenated aromatic hydrocarbons in naturally exposed European otters (*Lutra lutra*). Environmental Toxicology and Pharmacology 6, 91-102.

Van den Brink, N.W., Jansman, H.A.H. (2006). Applicability of spraints for monitoring organic contaminants in free-ranging otters (*Lutra lutra*). Environmental Toxicology & Chemistry 25, 2821-2826.

5.3.2. Question 1

Name three reasons why assessment of risk of PCBs to otters is relatively complicated

---

5.3.2. Question 2

How is it possible that although toxicity threshold levels are exceeded in some otter populations, for example in Scotland, the population seem to thrive really well?

This page titled 2.3.5.3: Wildlife population ecotoxicology is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **5.3: Wildlife population ecotoxicology** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.5.4: Trait-based approaches

5.4. Trait-based approaches

Author: Paul J. Van den Brink

Reviewers: Nico van den Brink, Michiel Kraak, Alexa Alexander-Trusiak

Learning objectives:

You should be able to

- describe how the characteristics (traits) of species determine their sensitivity, recovery and the propagation of effects to higher levels of biological organisation.
- explain the concept of response and effect traits.
- explain how traits-based approaches can be implemented into environmental risk assessment.

Keywords: Sensitivity, levels of biological organisation, species traits, recovery, indirect effects

Introduction

It is impossible to assess the sensitivity of all species to all chemicals. Risk assessments therefore, need methods to extrapolate the sensitivity of a limited number of species to all species present in the environment is desired. Statistical approaches, like the species sensitivity distribution concept, perform this extrapolation by fitting a statistical distribution (e.g. log-normal distribution) to a selected set of sensitivity data (e.g. 96h-EC50 data) in order to obtain a distribution of the sensitivity of all species. From this distribution a threshold value associated with the lower end of the distribution can be chosen and used as a protective threshold value (Figure 1).

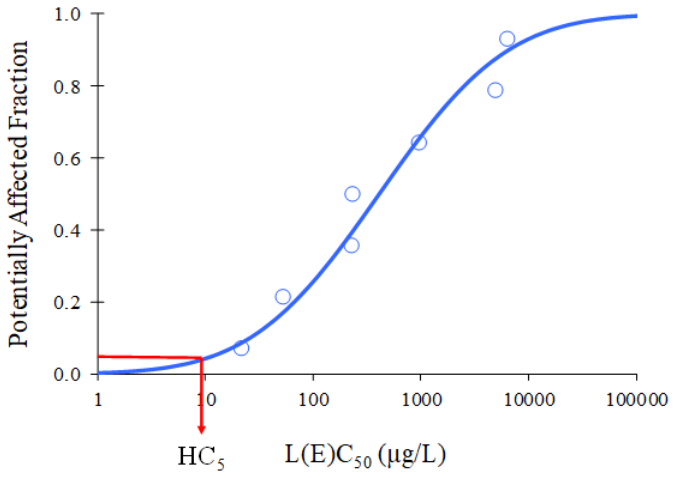


Figure 1. Species sensitivity distribution (line) fitted through a set of EC50 values (dots) and the threshold value protective for at least 95% of the species (Hazardous Concentration 5%, corresponding to the potentially affected fraction of species of 5%) derived from this distribution.

The disadvantage of this approach is that it does not include mechanistic knowledge on what determines species' sensitivity and uses species taxonomy rather than their characteristics. To overcome these and other problems associated with a taxonomy based approach (see Van den Brink et al., 2011 for a review) traits-based bioassessment approaches have been developed for assessing the effects of chemicals on aquatic ecosystem. In traits-based bioassessment approaches, species are not represented by their taxonomy but by their traits. A trait is a phenotypic or ecological characteristic of an organism, usually measured at the individual level but often applied as the average state/condition of a species. Examples of traits are body size, feeding habits, food preference, mode of respiration and lipid content. Traits describe the physical characteristics, ecological niche, and functional role of a species within the ecosystem. The recognized strengths of traits-based bioassessment approaches include: (1) traits add mechanistic and diagnostic knowledge, (2) traits are transferrable across geographies, (3) traits require no new sampling methodology as data that are currently collected can be used, (4) the use of traits has a long-standing tradition in ecology and can supplement taxonomic analysis.

When traits are used to study effects of chemical stressors on ecosystem structure (community composition) and function (e.g. nutrient cycling) it is important to make a distinction between response and effects traits (Figure 2). Response traits are traits that enable a response of the species to the exposure to a chemical. An example of a response trait may be size related surface area of an organism. Smaller organisms have relatively large surface areas, because their surface to volume ratio is higher than for larger animals Herewith, the uptake rate of the chemical stressor is generally higher in smaller animals compared to larger ones (Rubach et al., 2012). Effects traits of organisms influence the surrounding environment by the organisms, by altering the structure and functioning of the ecosystem. An example of an effect trait is the food preference of an organism. For instance, if the small (response trait) and herewith sensitive organisms happen to be herbivorous (effect trait) an increase in algal biomass may be expected when the organisms are affected (Van den Brink, 2008). So, to be able to predict ecosystem level responses it is important to know the (cor)relations between response and effect traits as traits are not independent from each other but can be linked phylogenetically or mechanistically and thus form trait syndromes (Van den Brink et al., 2011).

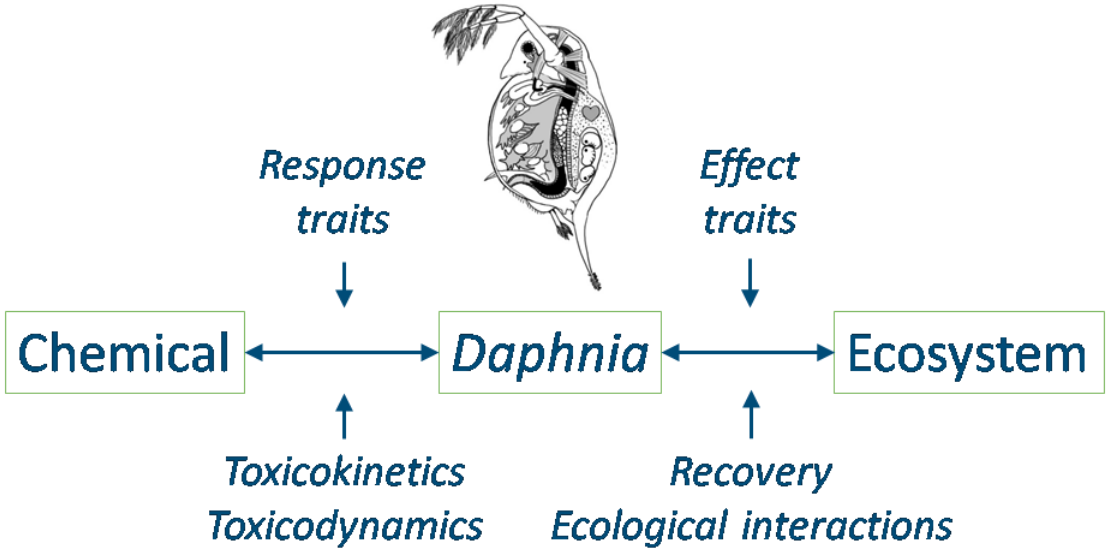


Figure 2.
Conceptual diagram showing how species' traits determine both the sensitivity of species to chemicals and how the effects propagate to higher levels of biological organisation.

Predictive models for sensitivity using response traits

One of the holy grails of ecotoxicology is to find out which species traits make one species more sensitive to a chemical stressor than another one. In the past, two approaches have been used to assess the (cor)relationships between species traits and their sensitivity, one based on empirical correlations between species' traits and their sensitivity as represented by EC50's (Rico and Van den Brink, 2015) and one based on a more mechanistic approach using toxicokinetic/toxicodynamic experiments and models (Rubach et al., 2012). Toxicokinetic-toxicodynamic models (TKTD models) simulate the time-course of processes leading to toxic effects on organisms (Jager et al., 2011). Toxicokinetics describe what an individual does with the chemical and, in their simplest form, include the processes of uptake and elimination, thereby translating an external concentration of a toxicant to an internal body concentration over time (see Section on Toxicokinetics and Bioaccumulation). Toxicodynamics describes what the chemical does with the organism, herewith linking the internal concentration to the effect at the level of the individual organism over time (e.g., mortality) (Jager et al., 2011) (see Sections on Toxicokinetics and Bioaccumulation and on Toxicodynamics and Molecular Interactions). Rubach et al. (2012) showed that almost 90% of the variation in uptake rates and 80% of the variation in elimination rates of an insecticide in a range of 15 freshwater arthropod species could be explained by 4 species traits. These traits were: i) surface area (without gills), ii) detritivorous feeding, iii) using atmospheric oxygen and iv) phylogeny in case of uptake, and i) thickness of exoskeleton, ii) complete sclerotization, iii) using dissolved oxygen and iv) % lipid of dry weight in case of elimination. For most of these traits, a mechanistic hypothesis between the traits and their influence on the uptake and elimination can be made (Rubach et al., 2012). For instance, a higher surface area to volume ratio increases the uptake of the chemical, so uptake is expected to be higher in small animals compared to larger animals. This shows that it is possible to construct mechanistic

models that are able to predict the toxicokinetics of chemicals in species and herewith the sensitivity of species to chemicals based on their traits.

The use of effect traits to model recovery and indirect effects

Traits determining the way organisms within an ecosystem react to chemical stress are related to the intrinsic sensitivity of the organisms on the one hand (response traits) and their recovery potential and food web relations (effect traits) on the other hand (Van den Brink, 2008). Recovery of aquatic invertebrates is, for instance, determined by traits like number of life cycles per year, the presence of insensitive life stages like resting eggs, dispersal ability and having an aerial life stage (Gergs et al., 2011) (Figure 3).

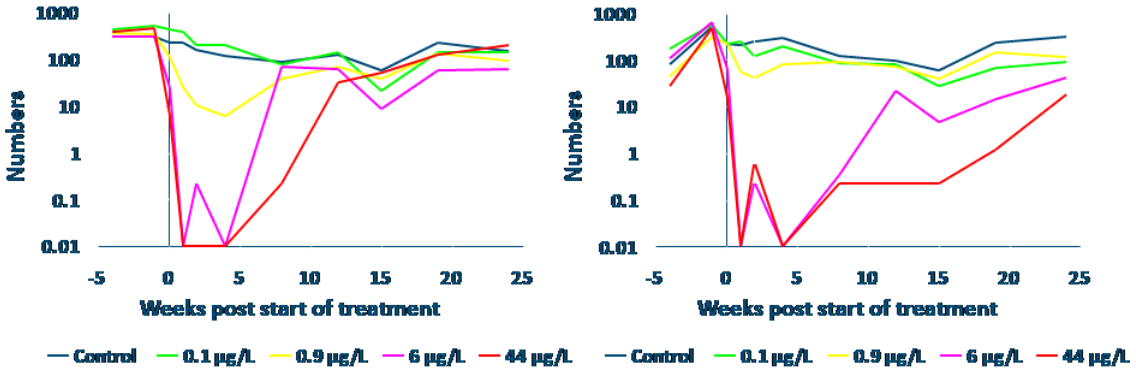


Figure 3.
Effect and recovery patterns as observed for two mayfly species after a pulsed exposure to an insecticide. Both mayflies are equally susceptible to chlorpyrifos, but the species on the left (Cloeon dipterum) has many life cycles per year while the species to the right (Caenis horaria) has a full life cycle in spring (coinciding with the chlorpyrifos treatment) and a partial one in autumn.

Besides recovery, effect traits will also determine how individual level effects will propagate to higher levels of biological organisation like the community or ecosystem level. For instance, when *Daphnia* are affected by a chemical, their trait related to food preference (algae) will ensure that, under nutrient-rich conditions, the algae will not be subjected to top-down control and will increase in abundance. The latter effects are called indirect effects, which are not a direct result of the exposure to the toxicant but an indirect one through competition, food-web relationships, etc..

References

Gergs, A., Classen, S., Hommen, U. (2011). Identification of realistic worst case aquatic macroinvertebrate species for prospective risk assessment using the trait concept. Environmental Science and Pollution Research 18, 1316-1323.

Jager, T., Albert, C., Preuss, T.G., Ashauer, R. (2011). General unified threshold model of survival-A toxicokinetic-toxicodynamic framework for ecotoxicology. Environmental Science and Technology 45, 2529-2540

Rico, A., Van den Brink, P.J. (2015). Evaluating aquatic invertebrate vulnerability to insecticides based on intrinsic sensitivity, biological traits and toxic mode-of-action. Environmental Toxicology and Chemistry 34, 1907-1917.

Rubach, M.N., D.J. Baird, M-C. Boerwinkel, S.J. Maund, I. Roessink, Van den Brink P.J. (2012). Species traits as predictors for intrinsic sensitivity of aquatic invertebrates to the insecticide chlorpyrifos. Ecotoxicology 21, 2088-2101.

Van den Brink, P.J. (2008). Ecological risk assessment: from book-keeping to chemical stress ecology. Environmental Science and Technology 42, 8999-9004.

Van den Brink P.J., Alexander, A., Desrosiers, M., Goedkoop, W., Goethals, P., Liess, M., Dyer, S. (2011). Traits-based approaches in bioassessment and ecological risk assessment: strengths, weaknesses, opportunities and threats. Integrated Environmental Assessment and Management 7, 198-208.

5.4. Question 1

Which two traits may determine the sensitivity of invertebrates to insecticides?

5.4. Question 2

Will two daphnids which are equal except in their size be equally sensitive to a chemical?

5.4. Question 3

What is the difference between a response and an effect trait?

5.4. Question 4

Which traits are of importance for the recovery of impacted populations?

5.4. Question 5

How will the insecticide-induced death of water fleas propagate to the ecosystem level under eutrophic circumstances, and what are these types of effects called?

This page titled 2.3.5.4: Trait-based approaches is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **5.4: Trait-based approaches** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.5.5: Population models

5.5. Population models

Authors: A. Jan Hendriks and Nico van Straalen

Reviewers: Aafke Schipper, John D. Stark and Thomas G. Preuss

Learning objectives

You should be able to

- explain the assumptions underlying exponential and logistic population modelling
- calculate intrinsic population growth rate from a given set of demographic data
- outline the conclusions that may be drawn from population modelling in ecotoxicology
- indicate the possible contribution of population models to chemical risk assessment

Keywords: intrinsic rate of increase, carrying capacity, exponential growth,

Introduction

Ecological risk assessment of toxicants usually focuses on the risks run by individuals, by comparing exposures with no-effect levels. However, in many cases it is not the protection of individual plants or animals that is of interest but the protection of a viable population of a species in an ecological context. Risk assessment generally does not take into account the quantitative dynamics of populations and communities. Yet, understanding and predicting effects of chemicals at levels beyond that of individuals is urgently needed for several reasons. First, we need to know whether quality standards are sufficiently but not overly protective at the population level, when extrapolated from toxicity tests. Second, responses of isolated, homogenous cohorts in the laboratory may be different from those of interacting, heterogeneous populations in the field. Third, to set the right priorities in management, we need to know the relative and cumulative effect of chemicals in relation to other environmental pressures.

Ecological population models for algae, macrophytes, aquatic invertebrates, insects, birds and mammals have been widely used to address the risk of potentially toxic chemicals, however, until recently, these models were only rarely used in the regulatory risk assessment process due to a lack of connection between model output and risk assessment needs (Schmolke et al., 2010). Here, we will sketch the basic principles of population dynamics for environmental toxicology applications.

Exponential growth

Ecological textbooks usually start their chapter on population ecology by introducing exponential and logistic growth. Consider a population of size N. If resources are unlimited, and the *per capita birth* (b) and *death rates* (d) are constant in a population closed to migration, the number of individuals added to the population per time unit (dN/dt) can be written as:

$\frac{dN}{dt} = (b-d)N(t)$ or $\frac{dN}{dt} = rN(t)$

where r is called the *intrinsic rate of increase*. This differential equation can be solved with boundary condition N(0) = N_0 to yield

$N(t) = N_0\, e^{rt}$ Figure 1a). It applies for example, to animal populations early in the growing season or when they have colonized a new environment. The global human population has also seen exponential growth during most of its existence. The tendency for any population to grow exponentially was recognized by Malthus in his book "An Essay on Population", published in 1789, and it helped Darwin to formulate his theory of evolution by natural selection.

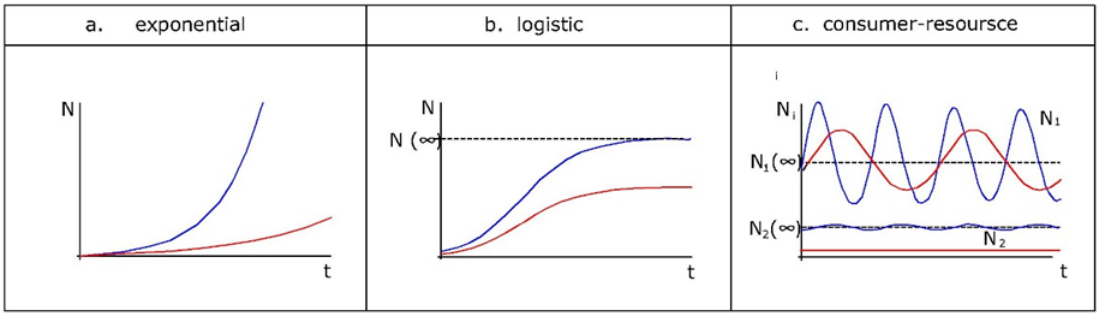


Figure 1. Population density as a function of time for exponential (a), logistic (b) and oscillating (c) populations without (blue) and with (red) exposure to toxicants. N = density, N(∞) equilibrium density, N_1 = resource density, N_2 = consumer density. Drawn by Wilma IJzerman.

Since toxicants will affect either reproduction or survival, or both, they will also affect the exponential growth rate (Figure 1a). This suggests that r can be considered a measure of population performance under toxic stress. But rather than from observed population trajectories, r is usually estimated from life-history data. We know from basic demographic theory that any organism with "time-invariant" vital rates (that is, fertility and survival may depend on age, but not on time), will be growing exponentially at rate r. The intrinsic rate of increase can be derived from age-specific survival and fertility rates using the so-called Euler-Lotka equation, which reads:

$$\int_0^{x_m} l(x)\, m(x)\, e^{-rx}\, dx = 1$$

in which x is age, x_m maximal age, l(x) survivorship from age zero to age x and m(x) the number of offspring produced per time unit at age x. Unfortunately this equation does not allow for a simple derivation of r; r must be obtained by iteration and the correct value is the one that, when combined with the l(x) and m(x) data, makes the integral equal to 1. Due to this complication approximate approaches are often applied. For example, in many cases a reasonably good estimate for r can be obtained from the age at first reproduction α, survival to first reproduction, S, and reproductive output, m, according to the following formula:

$$r = \frac{ln(S\, m)}{\alpha}$$

This is due to the fact that for many animals in the environment, especially those with high reproductive output and low juvenile survivorship, age at first reproduction is the dominant variable determining population growth (Forbes and Calow, 1999).

The classical demographic modelling approach, including the Euler-Lotka equation, considers time as a continuous variable and solves the equations by calculus. However, there is an equivalent formalism based on discrete time, in which population events are assumed to take place only at equidistant moments. The vital rates are then summarized in a so-called Leslie matrix, a table of survival and fertility scores for each age class, organized in such a way that when multiplied by the age distribution at any moment, the age distribution at the following time point is obtained. This type of modelling lends itself more easily to computer simulation. The outcome is much the same: if the Leslie matrix is time-invariant the population will grow each time step by a factor λ, which is related to r as ln λ = r (λ = 1 corresponds to r = 0). Mathematically speaking λ is the dominant eigenvalue of the Leslie matrix. The advantage of the discrete-time version is that λ can be more easily decomposed into its component parts, that is, the life-history traits that are affected by toxicants (Caswell, 1996).

The demographic approach to exponential growth has been applied numerous times in environmental toxicology, most often in studies of water fleas (Suhett et al., 2015), and insects (Stark and Banks, 2003). The tests are called "life-table response experiments" (see section on Population ecotoxicology in a laboratory setting). The investigator observes the effects of toxicants on age-specific survival and fertility, and calculates r as a measure of population performance for each exposure concentration. An example is given in Figure 2, derived from a study by Barata et al. (2000). Forbes and Calow (1999) concluded that the use of r in ecotoxicology adds ecological relevance to the analysis, but it does not necessarily provide a more sensitive or less sensitive endpoint: r is as sensitive as the vital rates underlying its estimation.

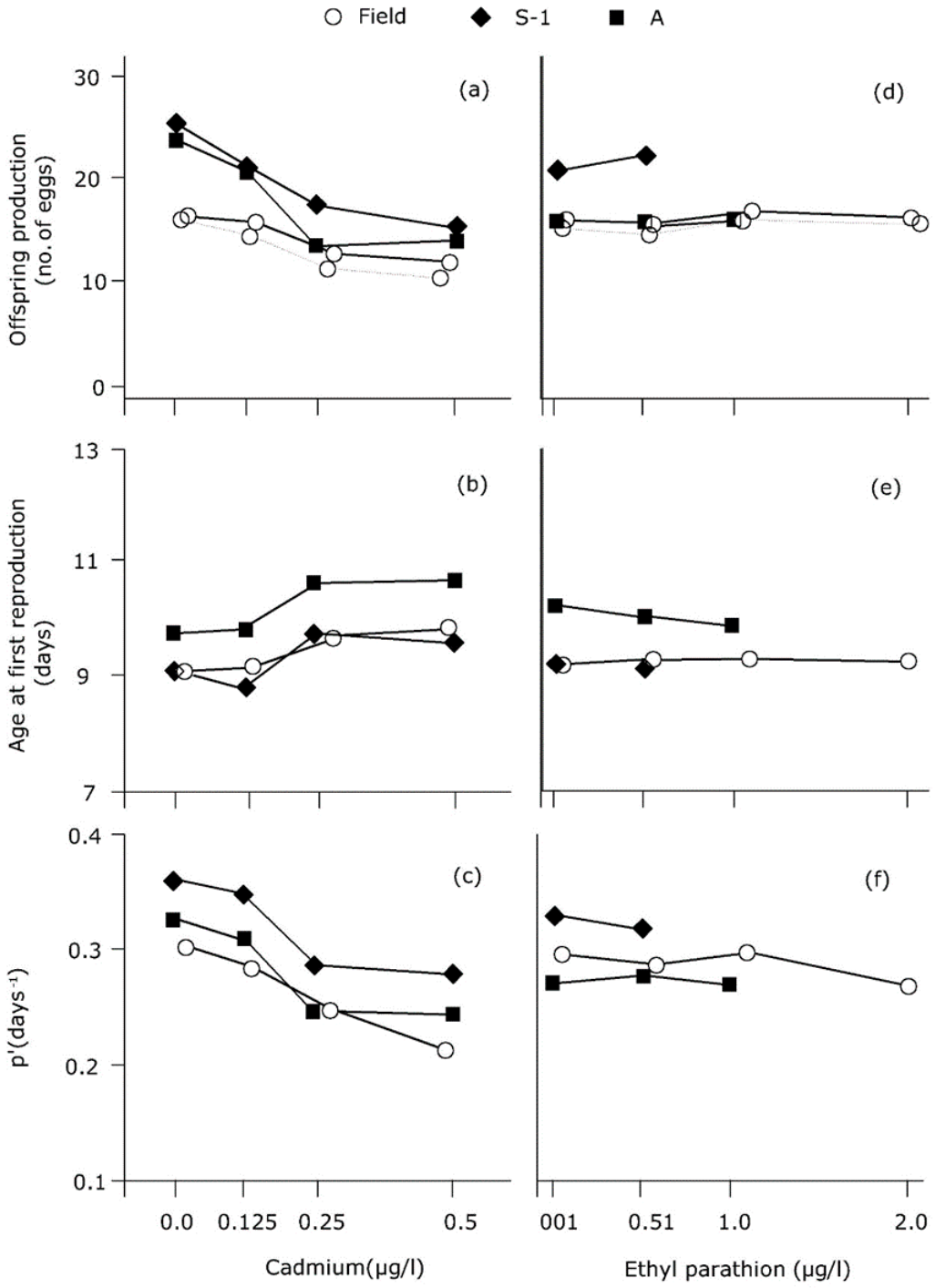


Figure 2. Example of results obtained in life-table response experiments with the water flea Daphnia magna. Cohorts of water fleas were cultured from field-collected resting eggs (ephippia) and two clonal lines under continuous culture in the laboratory (S-1 and A). Responses are shown for offspring production, age at first reproduction and the calculated intrinsic rate of population increase, as a function of cadmium and ethyl-parathion. Redrawn from Barata et al (2000) by Wilma IJzerman.

Hendriks et al. (2005) postulated that r should show a near-linear decrease with the concentration of a chemical, scaled to the LC_{50} (Figure 3). This relationship was confirmed in a meta-analysis of 200 laboratory experiments, mostly concerning invertebrate species (Figure 3). Anecdotal underpinning for large vertebrates comes from field cases where pollution limits population development.

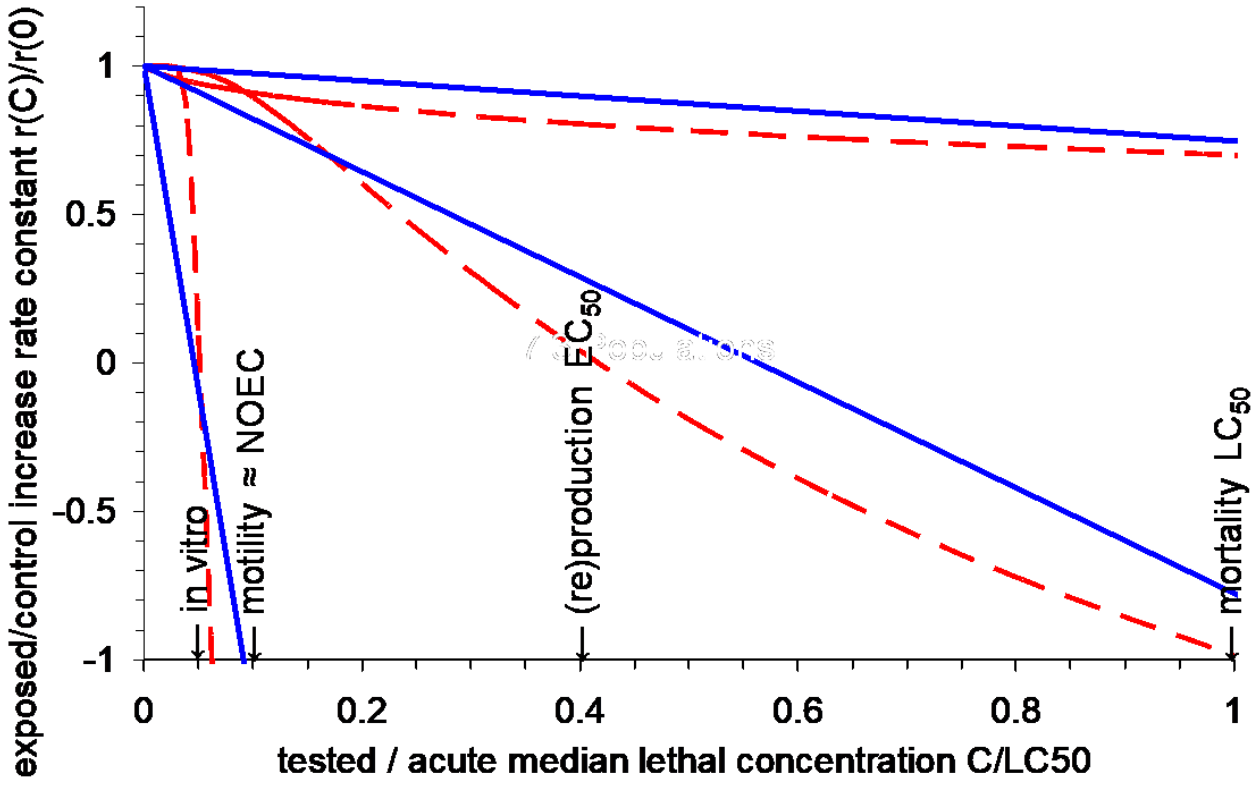


Figure 3. Modelled (red, μ ± 95% CI) and measured (blue, 5%-, 50%- and 95%-percentiles) population increase rates of organisms as a function of toxicant exposure concentrations C. Population growth r is expressed as the value observed in exposed organisms (r(C)) relative to the value in the control situation (r(0)) The exposure concentrations C, are scaled relative to the median lethal concentration LC_{50} of the toxicant. Characteristic (average) sub-lethal endpoints are given as a reference. The regression lines (individual data points not shown) are derived from a meta-analysis comprising 200 published laboratory experiments (adapted after Hendriks et al., 2005).

Logistic growth

As exponentially growing populations are obviously rare, models that include some form of density-dependence are more realistic. One common approach is to assume that the birth rate b decreases with density due to increasing scarcity of resources. The simplest assumption is a linear decrease with N, expressed as follows:

$\frac{dN}{dt} = rN(t)\,(1 - \frac{N(t)}{K})$Figure 1b). This is the density at which the population will reach equilibrium, also called carrying capacity: N(∞) = K. The resulting model is known as the logistic model, also called the Verhulst-Pearl equation, described in 1844 by François Verhulst and rediscovered in 1920 by Raymond Pearl. The solution of the logistic differential equation is a bit more complicated than in the case of simple exponential growth, but can be obtained by regular calculus to read: $N(t) = \frac{K}{1 + (\frac{K - N_0}{N_0})\,e^{-rt}}$ Figure 1b).

The question is, can the parameters of the logistic growth equation be used to measure population performance like in the case of exponential growth? Practical application is limited because the carrying capacity is difficult to measure under natural and contaminated conditions. Many field populations of arthropods, for example, fluctuate widely due to predator-prey dynamics, and hardly ever reach their carrying capacity within a growing season. An experimental study on the springtail *Folsomia candida* (Noël et al., 2006) showed that zinc in the diet did not affect the carrying capacity of contained laboratory populations, although there were several interactions below K that were influenced by zinc, including hormesis (growth stimulation by low doses of a toxicant), and Allee effects (loss of growth potential at low density due to lower encounter rate).

Density-dependence is expected to act as buffering mechanism at the population level because toxicity-induced population decline diminishes competition, however, the effects very much depend on the details of population regulation. This was demonstrated in a model for peregrine falcon exposed to DDE and PBDEs (Schipper et al., 2013). While the equilibrium size of the population

declined by toxic exposure, the probability of individual birds finding a suitable territory increased. However, at the same time the number of non-breeding birds shifting to the breeding stage became limiting and this resulted in a strong decrease in the equilibrium number of breeders.

Mechanistic effect models

To enhance the potential for application of population models in risk assessment, more ecological details of the species under consideration must be included, e.g. effects of dispersal, abiotic factors, predators and parasites, dispersal, landscape structure and many more. A further step is to track the physiology and ecology of each individual in the population. This is done in the dynamic energy budget modelling approach (DEB) developed by (Kooijman et al., 2009). By including such details, a model will become more realistic, and more precise predictions can be made on the effects of toxic exposures. These types of models are generally called "mechanistic effect models' (MEMs). They allow a causal link between the protection goal, a scenario of exposure to toxicants and the adverse population effects generated by model output (Hommen et al., 2015). The European Food Safety Authority (EFSA) in 2014 issued an opinion paper containing detailed guidelines on the development of such models and how to adjust them to be useful in the risk assessment of plant protection products.

References

Caswell, H. (1996). Demography meets ecotoxicology: untangling the population level effects of toxic substances. In: Newman, M.C., Jagoe, C.H. (Eds.). *Ecotoxicology. A hierarchical treatment.* Lewis Publishers**,** Boca Raton, pp. 255-292.

Barata, C., Baird, D.G., Amata, F., Soares, A.M.V.M. (2000). Comparing population response to contaminants between laboratory and field: an approach using *Daphnia magna* ephippial egg banks. Functional Ecology 14, 513-523.

EFSA (2014). Scientific Opinion on good modeling practice in the context of mechanistic effect models for risk assessment of plant protection products. EFSA Panel on Plant Protection and their Residues (PPR). EFSA Journal 12, 3589.

Forbes, V.E., Calow, P. (1999). Is the per capita rate of increase a good measure of population-level effects in ecotoxicology. Environmental Toxicology and Chemistry 18, 1544-1556.

Hendriks, A.J., Maas, J.L., Heugens, E.H.W., Van Straalen, N.M. (2005). Meta-analysis of intrinsic rates of increase and carrying capacity of populations affected by toxic and other stressors. Environmental Toxicology and Chemistry 24, 2267-2277

Hommen, U., Forbes, V., Grimm, V., Preuss, T.G., Thorbek, P., Ducrot, V. (2015). How to use mechanistic effect models in environmental risk assessment of pesticides: case studies and recommendations from the SETAC workshop Modelink. Integrated Environmental Assessment and Management 12, 21-31.

Kooijman, S.A.L.M., Baas, J., Bontje, D., Broerse, M., Van Gestel, C.A.M., Jager, T. (2009). Ecotoxicological Applications of Dynamic Energy Budget theory. In: Devillers, *J. (Ed.). Ecotoxicology Modeling*, *Volume* 2, Springer, Dordrecht, pp. 237-260.

Noël, H.L., Hopkin, S.P., Hutchinson, T.H., Williams, T.D., Sibly, R.M. (2006). Towards a population ecology of stressed environments: the effects of zinc on the springtail *Folsomia candida*. Journal of Applied Ecology 43, 325-332.

Schipper, A.M., Hendriks, H.W.M., Kaufmann, M.J., Hendriks, A.J., Huijbregts, M.A.J. (2013). Modelling interactions of toxicants and density dependence in wildlife populations. Journal of Applied Ecology 50, 1469-1478.

Schmolke, A., Thorbek, P., Chapman, P., Grimm, V. (2010) Ecological models and pesticide risk assessment: current modelling practice. Environmental Toxicology and Chemistry 29, 1006-1012.

Stark, J.D., Banks, J.E. (2003) Population effects of pesticides and other toxicants on arthropods. Annual Review of Entomology 48, 505-519.

Suhett, A.L. et al. (2015) An overview of the contribution of studies with cladocerans to environmental stress research. Acta Limnologica Brasiliensia 27, 145-159.

This page titled 2.3.5.5: Population models is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **5.5: Population models** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.5.6: Metapopulations

5.6. Metapopulations

Author: Nico van den Brink

Reviewers: Michiel Kraak, Heikki Setälä,

Learning objectives

You should be able to

- explain the relevance of meta-population dynamics for environmental risks of chemicals
- name the important mechanisms linking meta-populations to chemical risks

Implications of meta-population dynamics on risks of environmental chemicals

Populations can be defined as a group of organisms from the same species which live in a specific geographical area. These organisms interact and breed with each other. At a higher level, one can define meta-populations which can be described as a set of spatially separated populations which interact to a certain extent. The populations may function separately, but organisms can migrate between the populations. Generally the individual populations occur in more or less favourable habitat patches which may be separated by less favourable areas. However, in between populations, good habitats may also occur, where populations have not yet established , or the local populations may have gone extinct. Exchange between populations within a meta-population depends on i) the distances between the individual populations, ii) the quality of the habitat between the populations, e.g. the availability of so-called stepping stones, areas where organisms may survive for a while but which are too small or of too low habitat quality to support a local population and iii) the dispersal potential of the species. Due to the interactions between the different populations within a meta-population, chemicals may affect species at levels higher than the (local) population, also at non-contaminated sites.

An important effect of chemicals at meta-population scale is that local populations may act as a source or sink for other populations within the meta-population. When a chemical affects the survival of organisms in a local population, the local population densities decline. This may increase the immigration of organisms from neighbouring populations within the meta-population. Decrease of local densities would decrease emigration, resulting in a net-influx of organisms into the contaminated site. This is the case when organisms do not sense the contaminants, or that the contaminants do not alter the habitat quality for the organisms. In case the immigration rate at the delivering/source population to replace the populations is too high to replace the leaving organisms, population densities in neighbouring populations may decline, even at the non-contaminated **source** sites. Consequently, local populations at contaminated sites may act as a **sink** for other populations within the meta-population, so chemicals may have a much broader impact than just local.

On the contrary, when the local population is relatively small, or the chemical stress is not chronic, meta-population dynamics may also mitigate local chemical stress. Population level impacts of chemicals may be minimised by influx of organisms of neighbouring populations, potentially **recovering** the population densities prior to the chemical stress. Such recovery depends on the extent and duration of the chemical impact on the local populations and the capacity of the other populations to replenish the loss of the organisms in the affected population.

Meta-population dynamics may thus alter the extent to which contaminants may affect *local* populations _ through migration between populations. However, chemicals may affect the total carrying capacity of the meta-population as a whole. This can be illustrated by the modelling approach developed by Levins in the late 1960s (Levins 1969). A first assumption in this model is that not all patches that can potentially carry a local population are actually occupied, so let F be the fraction of occupied patches (1-F being the fraction not occupied). Populations have an average change of extinction being e (day^{-1} when calculating on a daily base), while non-occupied patches have a change of c of being populated (day^{-1}) from the populated patches. The daily change in numbers of occupied patches is therefore:

Formule

In this formula c*F*(1-F) equals the number of non-occupied patches that are being occupied from the occupied patches, while e * F equals the fraction of patches that go extinct during the day. This can be recalculated to a carrying capacity (CC) of

$$CC = 1 - \frac{e}{c}$$

while the growth rate (GR) of the meta-population can be calculated by

$$GR = c - e$$

In case chemicals increase extinction risk (*e*), or decrease the chance on establishment in a new patch (*c*) this will affect the CC (which will decrease because *e*/*c* will increase) as well as the GR (will decrease, may even go below 0). However, this model uses average coefficients, which may not be directly applicable to individual contaminated sites within a meta-population. More (complex) recent approaches include the possibility to use local-population specific parameters and even more, such model include stochasticity, increasing their environmental relevance.

Besides affecting populations directly in their habitats, chemicals may also affect the areas between habitat patches. This may affect the potential of organisms to migrate between patches. This may decrease the chances of organisms to repopulate non-occupied patches, i.e. decrease *c*, and as such both CC and GR. Hence, in a meta-population setting chemicals even in a non-preferred habitat may affect long term meta-population dynamics.

References

Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America 15, 237-240

5.6. Question 1

What are drivers of recovery of local populations that are affected by a chemical stressor in a meta-population setting?

5.6. Question 2

In what type of meta-population would a local population be less affected, with smaller number of local populations which are relatively large, or a setup with lot of small local populations?

This page titled 2.3.5.6: Metapopulations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **5.6: Metapopulations** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.5.7: Community ecotoxicology

5.7. Community ecotoxicology

5.7.1. Community Ecotoxicology: theory and concepts

Authors: Michiel Kraak and Ivo Roessink

Reviewers: Kees van Gestel, Nico van den Brink, Ralf B. Schäfer

Learning objectives:

You should be able to

- motivate the importance of studying ecotoxicology at the community level.
- define community ecotoxicology and to name specific phenomena at the community and ecosystem level.
- explain the indirect effects observed in community ecotoxicology.
- explain how communities can be studied and how data from experiments at the community level can be analyzed.
- interpret community ecotoxicity data and to explain related challenges.

Keywords: Community ecotoxicology, species interactions, indirect effects, mesocosm, ecosystem processes.

Introduction

The motivation to study ecotoxicological effects at the community level is that generally the targets of environmental protection are populations, communities and ecosystems. Consequently, when scaling up research from the molecular level, via cells, organs and individual organisms towards the population, community or even ecosystem level the ecological and societal relevance of the obtained data strongly increase (Figure 1). Yet, the difficulty of obtaining data increases, due to the increasing complexity, lower reproducibility and the increasing time needed to complete the research, which typically involves higher costs. Moreover, when effects are observed in the field it may be difficult to link these to specific chemicals and to identify the drivers of the observed effects. Not surprisingly, ecotoxicological effects at the community and ecosystem level are understudied.

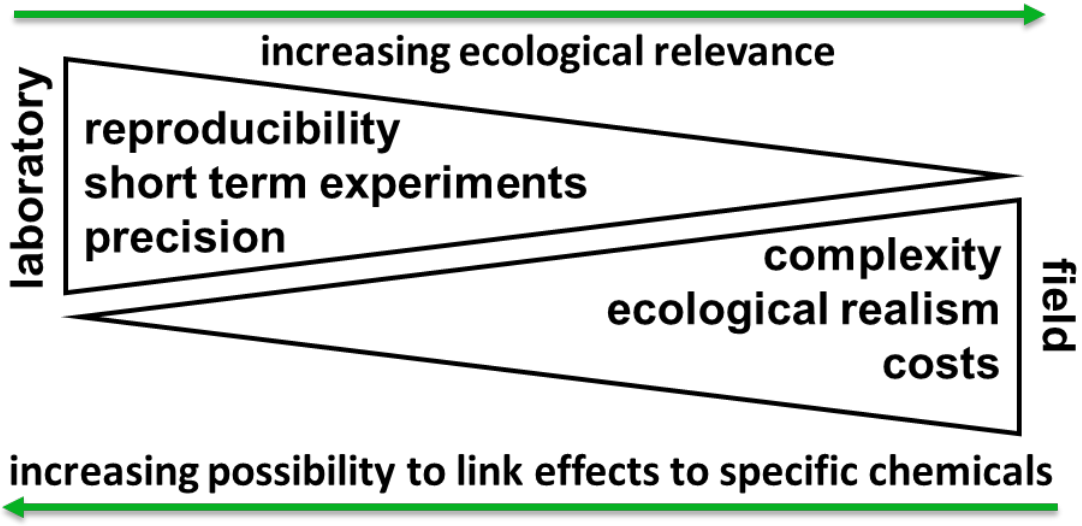


Figure 1.

Characteristics of ecotoxicological research performed at different levels of biological organization.

Community Ecotoxicology: definition and indirect effects

Community ecology is defined as the study of the organization and functioning of communities, which are assemblages of interacting populations of species living within a particular area or habitat. Building on this definition, community ecotoxicology is defined as the study of the effects of toxicants on patterns of species abundance, diversity, community composition and species interactions. These species interactions are unique to the community and ecosystem level and may cause direct effects of toxicants

on specific species to exert indirect effects on other species. It has been estimated that the majority of effects at these levels of biological organization are indirect rather than direct. These indirect effects are exerted via:

- predator-prey relationships
- consumer-producer relationships
- competition between species
- parasite-host relationships
- symbiosis
- biotic environment

As an example, Roessink et al. (2006) studied the impact of the fungicide triphenyltin acetate (TPT) on benthic communities in outdoor mesocosms. For several species a dose-related decrease in abundance directly after application was observed, followed by a gradual recovery coinciding with decreasing exposure concentrations, all implying direct effects of the fungicide. For some species, however, opposite results were obtained and abundance increased shortly after application, followed by a gradual decline; see the example of the Culicidae in Figure 2. In this case, these typical indirect effects were explained by a higher sensitivity of the predators and competitors of the Culicidae. Due to diminished predation and competition and higher food availability abundances of the Culicidae temporarily increased after toxicant exposure. With the decreasing exposure concentrations over time, the populations of the predators and competitors recovered, leading to a subsequent decline in numbers of the Culicidae.

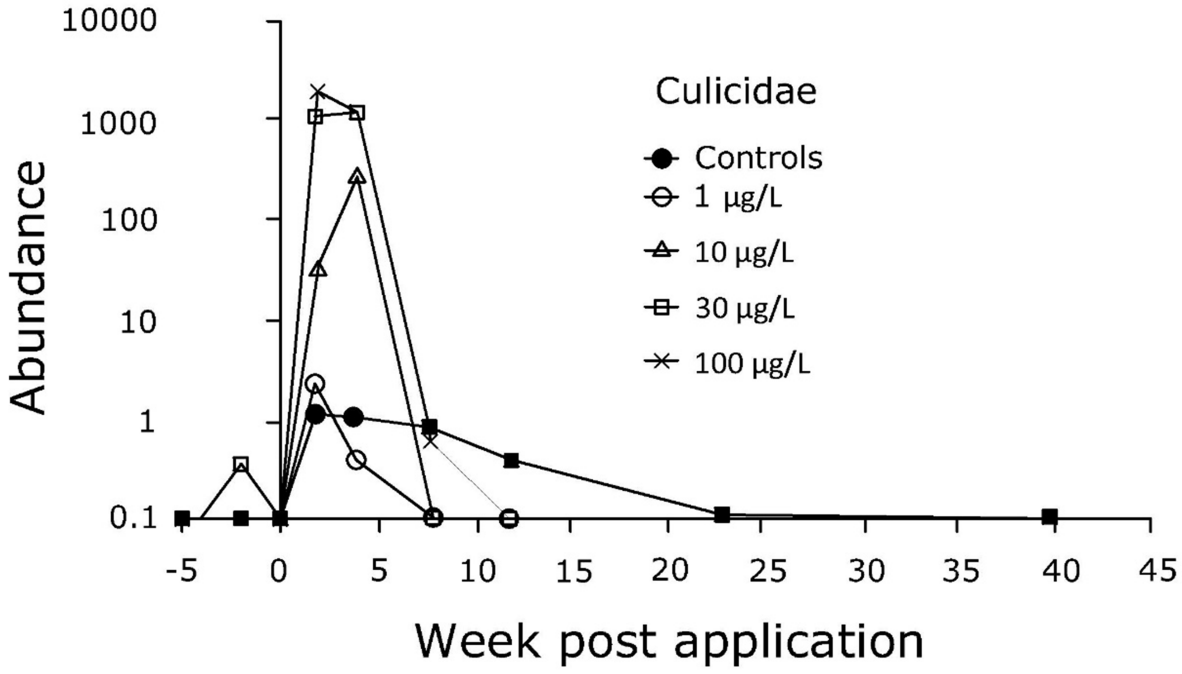


Figure 2. Dynamics of Culicidae in cosms treated with the fungicide triphenyltin acetate. Redrawn from Roessink et al. (2006) by Wilma IJzerman.

The indirect effects described above are thus due to species-specific sensitivities to the compound of interest, which influence the interactions between species. Yet, at higher exposure concentrations also the less sensitive species will start to be affected by the chemical. This may lead to an "arch-shaped" relationship between the number of individuals of a certain species and the concentration of a toxicant. In a mesocosm study with the insecticide lambda-cyhalothin this was observed for *Daphnia*, which are prey for the more sensitive phantom midge *Chaoborus* (Roessink et al., 2005; Figure 3). At low exposure concentrations the indirect effects, such as release from predation by *Chaoborus*, led to an increase in abundance of the less sensitive *Daphnia*. At intermediate exposure concentrations there was a balance between the positive indirect effects and the adverse direct effects of the toxicant. At higher exposure concentrations the adverse direct effects overruled the positive indirect effects resulting in a decline in abundance of the *Daphnia*. These combined dose dependent direct and indirect effects are inherent to community-level experiments, but are compound and species-interaction specific.

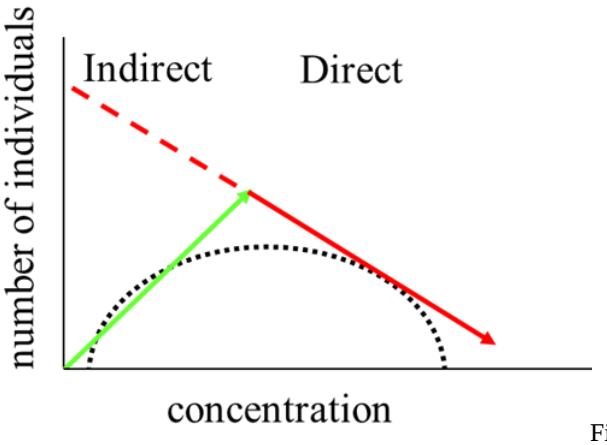


Figure 3. Arch-shaped relationship between the number of individuals of a certain species in a community-level experiment and the concentration of a toxicant, caused by the combination of indirect positive effects at low exposure concentrations and adverse direct effects at higher exposure concentrations.

Investigating communities and analysing and interpreting community ecotoxicity data

To study community ecotoxicology, experiments have to be scaled up and are therefore often performed in mesocosms, artificial ponds, ditches and streams, or even in the field, sometimes accompanied by the use of in- and exclosures. To assess the effects of toxicants on communities in such large systems requires meticulous sampling schemes, which often make use of artificial substrates and e.g. emergence traps for aquatic invertebrates with terrestrial adult life stages (see section on Community ecotoxicology in practice).

Alternatively to scaling up the experiments in community ecotoxicology, the size of the communities may be scaled down. Algae and bacteria grown on coin sized artificial substrates in the field or in experimental settings provide the unique advantage that the experimental unit is actually an entire community.

Given the large scale and complexity of experiments at the community level, such experiments generally generate overwhelming amounts of data, making appropriate analysis of the results challenging. Data analysis focusing on individual responses, so-called univariate analysis, that suffice in single species experiments, obviously falls short in community ecotoxicology, where cosm or (semi-)field communities sometimes consist of over a hundred different species. Hence, multivariate analysis is often more appropriate, similar to the approaches frequently applied in field studies to identify possible drivers of patterns in species abundances. Alternative approaches are also applied, like using ecological indices such as species richness or categorizing the responses of communities into effect classes (Figure 4). To determine if species under semi-field conditions respond equally sensitive to toxicant exposure as in the laboratory, the construction and subsequent comparison of species sensitivity distributions (SSD) (see section on SSDs) may be helpful.

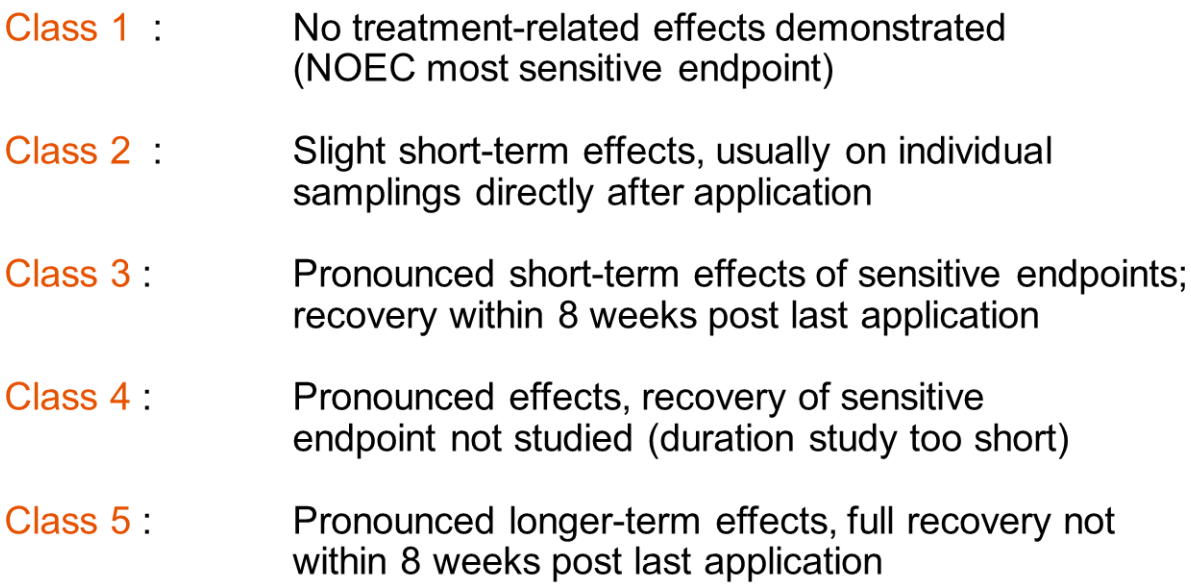


Figure 4. Effect classes to categorize the responses of communities in community-level experiments.

The analysis and interpretation of community ecotoxicity data is also challenged by the dynamic development of each individual replicate cosm, artificial pond, ditch or stream, including those from the control. From the start of the experiment, each control replicate develops independently, matures, and at the end of the experiments that generally last for several months control replicates may differ not only from the treatments, but also among each other. The challenge is then to separate the toxic signal from the natural variability in the data.

In experiments that include a recovery phase, it is frequently observed that previously exposed communities do recover, but develop in another direction than the controls, which actually challenges the definition of recovery. Moreover, recovery can be decelerated or accelerated depending on the dispersal capacity of the species potentially inhabiting the cosms and the distance to nearby populations within a metapopulation (see section on Metapopulations). Other crucial factors that may affect the impact of a toxicant on communities, as well as their recovery from this toxicant exposure include habitat heterogeneity and the state of the community in combination with the moment of exposure. Habitat heterogeneity may affect the distribution of toxicants over the different environmental compartments and may provide shelter to organisms. Communities generally exhibit temporal dynamics in species composition and in their contribution to ecosystem processes (see section on Structure versus function), as well in the lifecycle stages of the individual species. Exponentially growing populations recover much faster than populations that reached carrying capacity and for almost all species, young individuals are up to several orders of magnitude more sensitive than adults or late instar larvae (see section on Population ecotoxicology). Hence, the timing of exposure to toxicants may seriously affect the extent of the adverse effects, as well as the recovery potential of the exposed communities.

From community ecotoxicology towards ecosystems and landscapes

When scaling up from the community to the ecosystem level, again unique characteristics emerge: structural characteristics like biodiversity, but also ecosystem processes, quantified by functional endpoints like primary production, ecosystem respiration, nutrient cycling and decomposition. Although a good environmental quality is based on both ecosystem structure and functioning, there is definitely a bias towards ecosystem structure, both in science and in policy (see section on Structure versus function). Levels of biological organisation higher than ecosystems are covered by the field of landscape ecotoxicology (see section on Landscape ecotoxicology) and in a more practical way by the concept of ecosystem services (see section on Ecosystem services).

References

Roessink, I., Crum, S.J.H., Bransen, F., Van Leeuwen, E., Van Kerkum, F., Koelmans, A.A., Brock, T.C.M. (2006). Impact of triphenyltin acetate in microcosms simulating floodplain lakes. I. Influence of sediment quality. Ecotoxicology 15, 267-293.

Roessink, I., Arts, G.H.P., Belgers, J.D.M., Bransen, F., Maund, S.J., Brock, T.C.M. (2005). Effects of lambda-cyhalothrin in two ditch mesocosm systems of different trophic status. Environmental Toxicology and Chemistry 24, 1684-1696.

Further reading

Clements, W.H., Newman, M.C. (2002). Community Ecotoxicology. John Wiley & Sons, Ltd.

5.7.1. Question 1

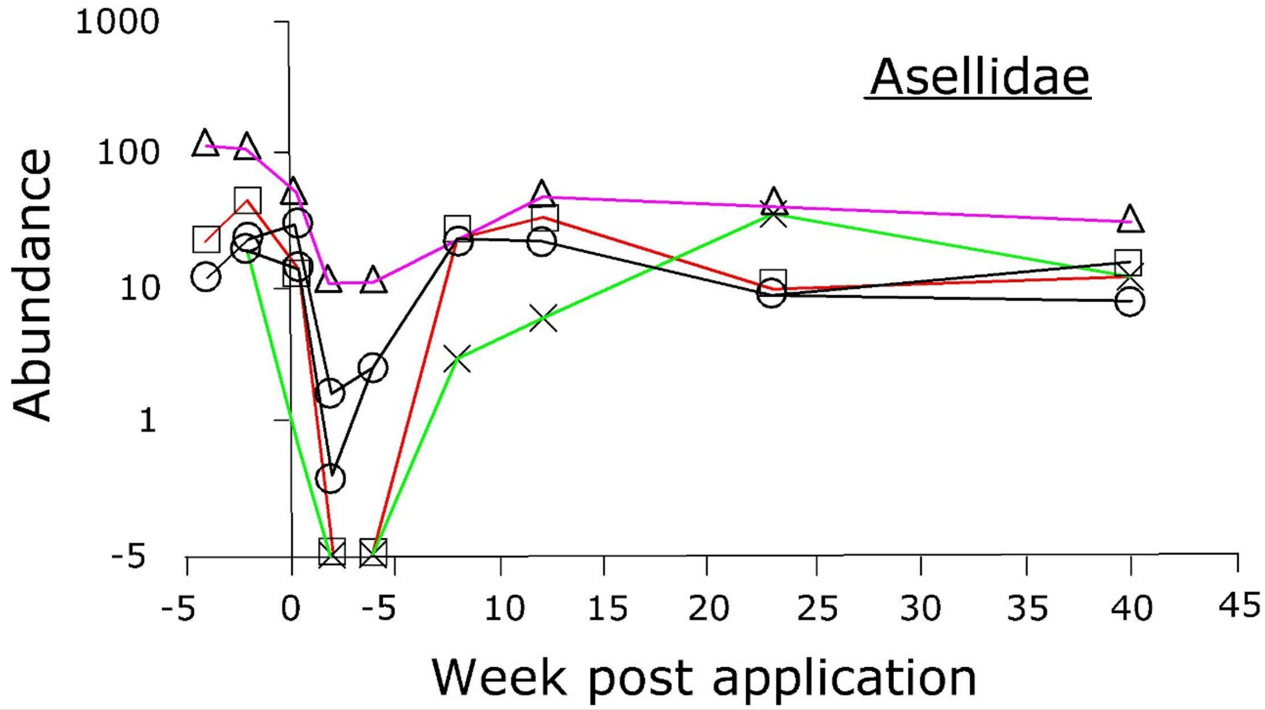
Motivate the importance of studying ecotoxicology at the community level.

5.7.1. Question 2

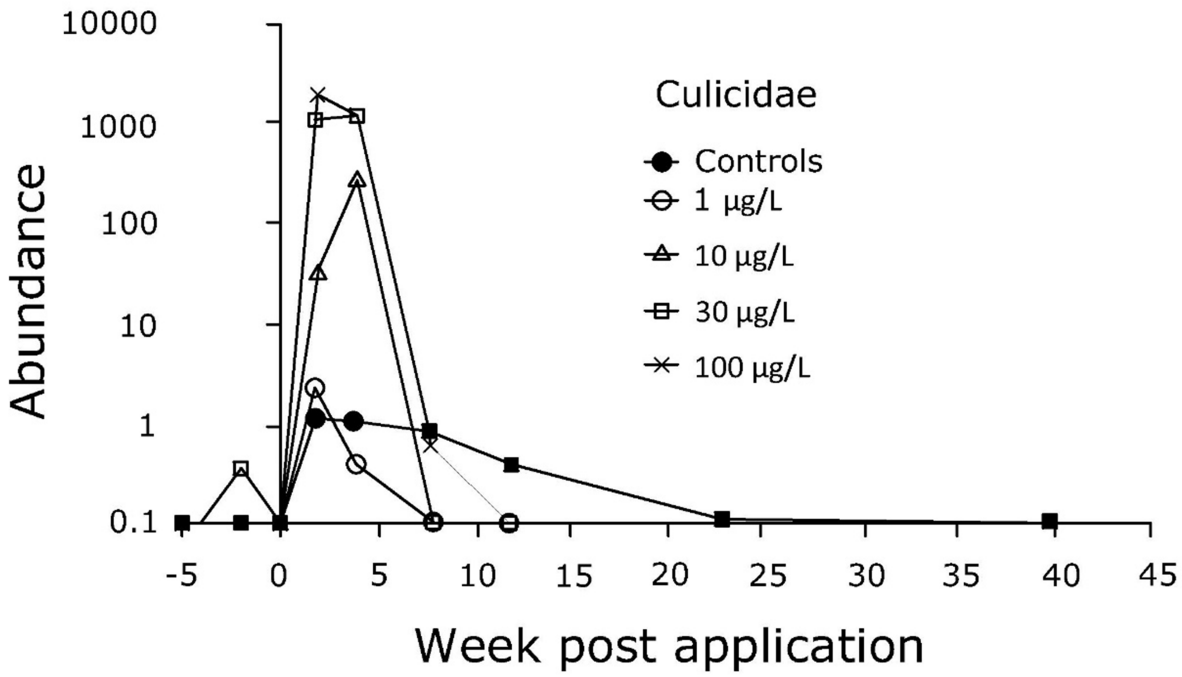
Define community ecotoxicology and name specific phenomena at the community and ecosystem level.

5.7.1. Question 3

Roessink et al. (2006) studied the impact of the fungicide triphenyltin acetate (TPT) on benthic communities in outdoor mesocosms. For several species they observed a dose related decrease in abundance directly after application, followed by a gradual recovery, see example in the graph below.



For some species however, opposite results were obtained and the abundance increased shortly after application, followed by a gradual decline, see example in the graph below.



Explain the results shown in the lower graph.

5.7.1. Question 4

Mention three ways to analyze data from experiments at the community level.

5.7.1. Question 5

Mention one advantage and three disadvantages of cosm experiments.

5.7.2. Community ecotoxicology in practice

Author: Martina G. Vijver

Reviewers: Paul J. van den Brink, Kees van Gestel

Learning objectives:

To be able to

- describe the variety of ecotoxicological test systems available to address different research questions.
- explain what type of information is gained from low as well as higher level ecotoxicological tests.
- explain the advantages and disadvantages of different higher level ecotoxicological test systems

Keywords: microcosms, mesocosms, realism, different biological levels

Introduction: Linking effects at different levels of biological organization

It is generally anticipated that ecotoxicological tests should provide data useful for making realistic predictions of the fate and effects of chemicals in natural ecosystems (Landner et al., 1989). The ecotoxicological test, if used in an appropriate way, should be able to identify the potential environmental impact of a chemical before it has caused any damage to the ecosystem. In spite of the considerable amount of work devoted to this problem and the plethora of test methods being published, there is still reason to question whether current procedures for testing and assessing the hazard of chemicals in the environment do answer the questions we have asked. Most biologists agree that at each succeeding level of biological organization new properties appear that would not have been evident even by the most intense and careful examination of lower levels of organization (Cairns Jr., 1983).

These levels of biological hierarchy might be crudely characterized as subcellular, cellular, organ, organism, population, multispecies, community, and ecosystem (Figure 1). At the lower biological level, responses are faster than those occurring at higher levels of organization.

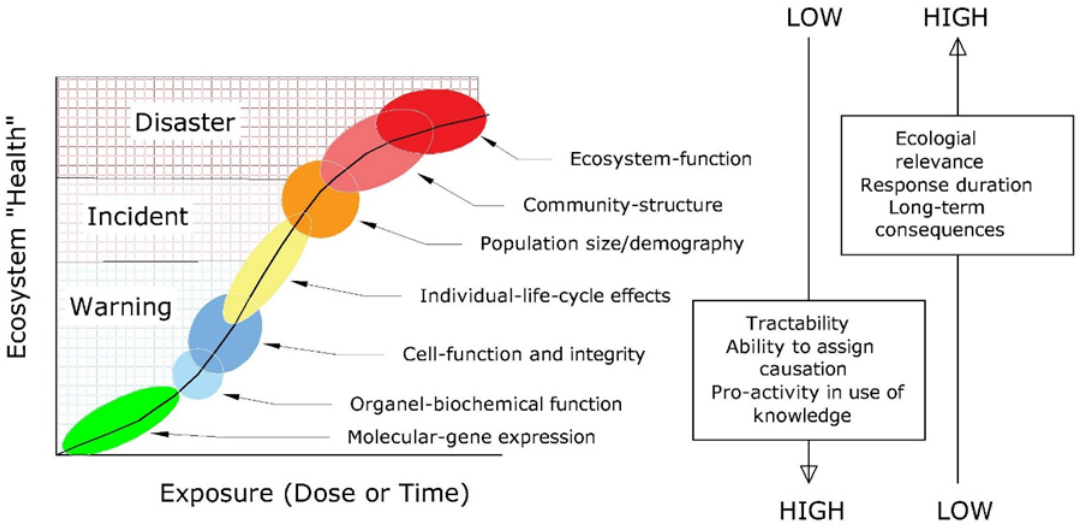


Figure 1: Tests at different biological levels, from molecule to ecosystem scale (modified from Newman, 2008). Each biological level is of equal importance in environmental toxicology, but has a different implication to the ecosystem health. If the fitness at the gene to the individual species level is affected due to exposure, this can be seen as a warning for the ecosystem health. If the impact of exposure can be detected at individual species (e.g. reproductive output) to population levels, this can be seen as an incident. If the impact of exposure is detectable at the community (structure and functioning) level, this is considered a disaster for ecosystem health. Measurements performed at the different biological level inform us differently: at the higher biological levels in general more ecological realism regarding exposure as well as species-interactions is gained, whereas at the lower biological levels causality and tractability of the link between the response and the dose of chemicals is achieved. Drawn by Wilma IJzerman.

Experiments executed at the lower biological level often are performed under standard laboratory conditions (see Section on Toxicity testing). The laboratory setting has advantages like allowing for replication, the use of relatively easy and simplified conditions that enable outcomes that are rather robust across different laboratories, the stressor of interest being more traceable under optimal stable conditions, and easy repetition of experiments. As a consequence, at the lower biological level the responses of organisms to chemical stressors tend to be more tractable, or more causal, than those identified when studying effects at higher tiered levels.

The merit to perform cosm studies, so at the higher biological level (see Figure 1), is to investigate the impact of a stressor on a variety of species, all having interactions with each other. This enables detecting both direct and indirect effects on the structure of species assemblages due to the chemicals. Indirect effects can become manifest as disruptions of species interactions, e.g. competition, predator-prey interactions and the like. A second important reason for conducting cosm studies is that abiotic interactions at the level of the ecosystem can be accounted for, allowing for measurement of effects of chemicals under more environmentally realistic exposure conditions. Conditions that likely influence the fate and behavior of chemical are sorption to sediments and plants, photolysis, changes in pH (see section on Bioavailability for a more detailed description), and other natural fluctuations.

What are cosm studies?

Microcosm or mesocosm (or cosm) studies represent a bridge between the laboratory and the natural world (examples of aquatic cosms are given in Figure 2). The difference between micro- and mesocosms is mostly restricted to size (Cooper and Barmuta, 1993). Aquatic microcosms are 10^{-3} to 10 m^3 in size, while mesocosms are 10 to 10^4 m^3 or even larger equivalent to whole or natural systems. The originality of cosms is mainly based on the combination of ecological realism, achieved by the introduction of the basic components of natural ecosystems, and facilitated access to a number of physicochemical, biological, and toxicological parameters that can be controlled to some extent. The cosm approach also makes it possible to work with treatments that can be replicated, so enabling the study of multiple environmental factors which can be manipulated. The system allows the establishment

of food webs, the assessment of direct and indirect effects, and the evaluation of effects of contamination on multiple trophic and taxonomic levels in an ecologically relevant context. Cosm studies make it possible to assess effects of contaminants by looking at the parts (individuals, populations, communities) and the whole (ecosystems) simultaneously.



Figure 2. Different aquatic ecotoxicological testing facilities: A) indoor microcosms (water-sediment interface, at Leiden University), B) in situ (or caged) outdoor enclosures (at Wageningen Environmental Research), and C) cosms or experimental ditches (at Living Lab, Leiden University).

As given in the OECD guidance document (OECD, 2004), the size to be selected for a meso- or microcosm study will depend on the objectives of the study and the type of ecosystem that is to be simulated. In general, studies in smaller systems are more suitable for short-term studies of up to three to six months and studies with smaller organisms (e.g. planktonic species). Larger systems are more appropriate for long-term studies (e.g. 6 months or longer). Numerous ecosystem-level manipulations have been conducted since the early 1970s (Hurlbert et al., 1972). The Experimental Lakes Area (ELA) situated in Ontario, Canada deserves special attention because of its significant contributions to the understanding of how natural communities respond to chemical stressors. This ELA consists of 46 natural, relatively undisturbed lakes, which were designated specially for ecosystem-level research. Many different questions have been tackled here, e.g. manipulations with nutrients (amongst others Levine and Schindler, 1999), synthetic estrogens (e.g. Kidd et al., 2014) and Wallace with pesticides in the Coweeta district (Wallace et al., 1999). It is nowadays realized that there is a need for testing more than just individual species and to take into account ecosystem elements such as fluctuations of abiotic conditions and biotic interactions when trying to understand the ecological effects of chemicals. Therefore a selection of study parameters is often considered as given by OECD (2004):

- Regarding *treatment regime*:
  - dosing regime, duration, frequency, loading rates, preparation of application

- solutions, application of test substance, etc.;
- meteorological records for outdoor cosms;
- physicochemical water parameters (temperature, oxygen saturation, pH, etc.);

- Regarding *biological levels* it should be recorded what sampling methods and taxonomic identification methods are used;
  - phytoplankton: chlorophyll-a; total cell density; abundance of individual dominant taxa; taxa (preferably species) richness, biomass;
  - periphyton: chlorophyll-a; total cell density; density of dominant species; species richness, biomass;
  - zooplankton: total density per unit volume; total density of dominant orders (Cladocera, Rotifera and Copepoda); species abundance; taxa richness, biomass;
  - macrophytes: biomass, species composition and % surface covering of individual plants;
  - emergent insects: total number emerging per unit of time; abundance of individual dominant taxa; taxa richness; biomass; density; life stages;
  - benthic macroinvertebrates: total density per unit area; species richness, abundance of individual dominant species; life stages;
  - fish: total biomass at test termination; individual fish weights and lengths for adults or marked juveniles; condition index; general behaviour; gross pathology; fecundity, if necessary.

Two typical examples of results obtained in an aquatic cosm study

A cosm approach assist in identifying and quantifying direct as well as indirect effects. Here two different types of responses are described, for more examples it is referred to the Section on Multistress.

Joint interactions: Barmentlo et al. (2018) used an outdoor mesocosm system consisting of 65 L ponds. Using a full factorial design, they investigated the population responses of macroinvertebrate species assemblages exposed for 35 days to environmentally relevant concentrations of three commonly used agrochemicals (imidacloprid, terbuthylazine, and NPK fertilizers). A detrivorous food chain as well as an algal-driven food chain were inoculated into the cosms. At environmentally realistic concentrations of binary mixtures, the species responses could be predicted based on concentration addition (see Section on Mixture toxicity). Overall, the effects of trinary mixtures were much more variable and counterintuitive. This was nicely illustrated by how the mayfly *Cloeon dipterum* reacted to the various combinations of the pesticides. Compared to single substance exposures and binary mixtures, extreme low recovery of *C. dipterum* (3.6% of control recovery for both mixtures) was seen. However, after exposure to the trinary mixture, recovery of *C. dipterum* no longer deviated from the control, and therefore was was higher than expected. Unexpected effects of the mixtures were also obtained for both zooplankton species (*Daphnia magna* and *Cyclops* sp.) As expected, the abundance of both zooplankton species was positively affected by nutrient applications, but pesticide addition did not lower their recovery. These type of unexpected results can only been identified when multiple species and multiple stressors are tested and cannot be detected in a lab-test with single species.

Indirect cascading effects: Van den Brink et al. (2009) studied the effects of chronic applications of a mixture of the herbicide atrazine and the insecticide lindane in indoor freshwater plankton-dominated microcosms. Both top-down and bottom-up regulation mechanisms of the species assemblage selected were affected by the pesticide mixture. Lindane exposure also caused a decrease in sensitive detritivorous macro-arthropods and herbivore arthropods. This allowed insensitive food competitors like worms, rotifers and snails to increase in abundance (although not always significantly). Atrazine inhibited algal growth and hence also affected the herbivores. A direct result of the inhibition of photosynthesis by atrazine exposure were lower dissolved oxygen and pH levels and an increase in alkalinity, nitrogen and electrical conductivity. See Figure 3 for a synthesis of all interactions observed in the study of Van den Brink et al. (2009).

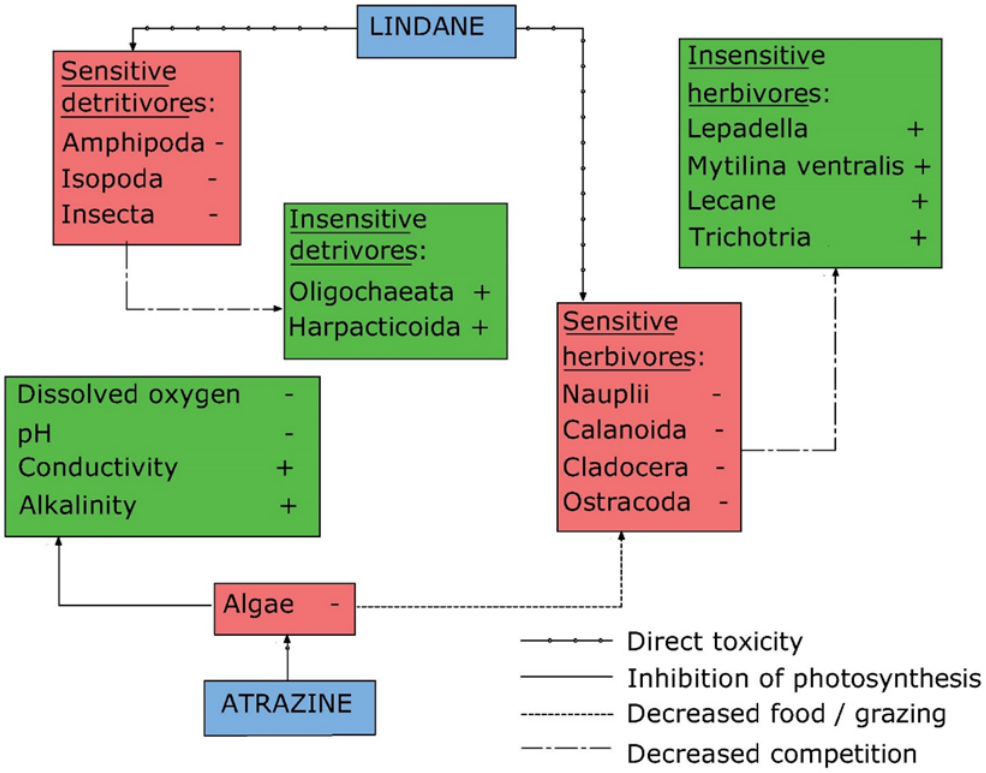


Figure 3. Ecological effect chain as observed in a microcosm experiments using atrazine and lindane as stressors. The arrows indicate the hypothesis-driven relationships between species. The red colors (with -) represent negative feedbacks, the green colors (with +) positive feedbacks. See the text for further explanation. Adapted from Van den Brink et al. (2009) by Wilma IJzerman.

Realism of cosm studies

There is a conceptual conflict between realism and replicability when applied to mesocosms. Replicability may be achieved, in part, by a relative simplification of the system. The crucial point in designing a model system may not be to maximize the realism, but rather to make sure that ecologically relevant information can be obtained. Reliability of information on ecotoxicological effects of chemicals tested in mesocosms closely depends on the representativeness of biological processes or structures that are likely to be affected. This means that within cosms key features at both structural and functional levels should be preserved as they ensure ecological representativeness. Extrapolation from small experimental systems to the real world seems generally more problematic than the use of larger systems in which more complex interactions can be studied experimentally as well. For that reason, Caquet et al. (2000) claim that testing chemicals using mesocosms refines the classical methods of ecotoxicological risk assessment because they provide conditions for a better understanding of environmentally relevant effects of chemicals.

References

Barmentlo S.H., Schrama M., Hunting E.R., Heutink R., Van Bodegom P.M., De Snoo G.R., Vijver M.G. (2018). Assessing combined impacts of agrochemicals: Aquatic macroinvertebrate population responses in outdoor mesocosms, Science of the Total Environment 631-632, 341-347.

Caquet, T., Lagadic, L., Sheffield, S.R. (2000) Mesocosm in ecotoxicology: outdoor aquatic systems. Reviews of Environmental Contamination and Toxicology 165, 1-38.

Cairns Jr. J. (1983). Are single species toxicity tests alone adequate for estimating environmental hazard? Hydrobiologica 100, 47-57.

Cooper, S.D., Barmuta, L.A. (1993) Field experiments in biomonitoring. In Rosenberg, D.M., Resh, V.H. (Eds.) Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman and Hall, New York, pp. 399-441.

OECD (2004). Draft Guidance Document on Simulated Freshwater Lentic Field Tests (Outdoor Microcosms and Mesocosms) (July 2004). Organization for Economic Cooperation and Development, Paris. http://www.oecd.org/fr/securitechimique/essais/32612239.pdf

Hurlbert, S.H., Mulla, M.S., Willson, H.R. (1972) Effects of an organophosphorus insecticide on the phytoplankton, zooplankton, and insect populations of fresh-water ponds. Ecological Monographs 42, 269-299.

Kidd, K.A., Paterson, M.J., Rennie, M.D., Podemski, C.L., Findlay, D.L., Blanchfield, P.J., Liber, K. (2014). Direct and indirect responses of a freshwater food web to a potent synthetic oestrogen. Philosophical Transactions of the Royal Society B Biological Sciences 369, Article AR 20130578, DOI:10.1098/rstb.2013.0578

Landner, L., Blanck, H., Heyman, U., Lundgren, A., Notini, M., Rosemarin, A., Sundelin, B. (1989) Community Testing, Microcosm and Mesocosm Experiments: Ecotoxicological Tools with High Ecological Realism. Chemicals in the Aquatic Environment. Springer, pp. 216-254.

Levine, S.N., Schindler, D.W. (1999). Influence of nitrogen to phosphorus supply ratios and physicochemical conditions on cyanobacteria and phytoplankton species composition in the Experimental Lakes Area, Canada. Canadian Journal of Fisheries and Aquatic Sciences 56, 451-466.

Newman, M.C. (2008). Ecotoxicology: The History and Present Directions. In Jørgensen, S.E., Fath, B.D. (Eds.), Ecotoxicology. Vol. 2 of Encyclopedia of Ecology, 5 vols. Oxford: Elsevier, pp.1195-1201.

Van den Brink, P.J., Crum, S.J.H., Gylstra, R., Bransen, F., Cuppen, J.G.M., Brock, (2009). Effects of a herbicide - insecticide mixture in freshwater microcosms: risk assessment and ecological effect chain. Environmental Pollution 157, 237-249.

Wallace, J.B., Grubaugh, J.W., Whiles, M.R. (1996). Biotic indices and stream ecosystem processes: Results from an experimental study. Ecological Applications 6, 140-151.

5.7.2. Question 1

Responses of organisms to long-term exposure can be detected at the sub-organism level by making use of biomarkers and then extrapolating these results to organism fitness and consequence at the population level. Mention two benefits of performing tests making use of biomarkers

---

5.7.2. Question 2

Mention at least two benefits of performing tests at the higher biological level such as community or ecosystem levels.

This page titled 2.3.5.7: Community ecotoxicology is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **5.7: Community ecotoxicology** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.5.8: Structure versus function incl. ecosystem services

5.8. Structure versus function incl. ecosystem services

Author: Herman Eijsackers

Reviewers: Nico van den Brink, Kees van Gestel, Lorraine Maltby

Learning objectives:

You should be able to

- mention three levels of biodiversity
- describe the difference between structural and functional properties of an ecosystem
- explain why the functioning of an ecosystem generally tends to be less sensitive than its structure
- describe the term Functional Redundancy and explain its meaning for interpreting effects on the structure and functioning of ecosystems

Keywords: structural biodiversity, functional biodiversity, functional redundancy, food web interactions

Biodiversity at three different levels

In ecology, biodiversity describes the richness of natural life at three levels: *genetic diversity*, *species diversity* (the most well-known) and *landscape diversity*. The most commonly used index, the Shannon Wiener index, expresses biodiversity in general terms as the number of species in relation to the number of individuals per species. Precisely, this index stands for the sum of the natural logarithm of the number of individuals per species present:

$-\sum(p_i*\ln(p_i))$

with $p_i = n_i/N$ in which n_i is the number of individuals of species i and N the total number of individuals of all species combined.

In environmental toxicology, most attention is paid to species diversity. Genetic diversity plays a role in the assessment of more or less sensitive or resistant subspecies or local populations of a species, like in various mining areas. Landscape diversity is receiving attention only recently and aims primarily at the total load of e.g. pesticides applied in an agronomic landscape (see Section on Landscape ecotoxicology), although it should more logically focus on the interactions between the various ecosystems in a landscape, for instance a lake surrounded partly by a forest, partly by a grassland.

Structural and functional diversity

In general, the various types of interactions between species do not play a major role in the study of biodiversity neither within ecology nor in environmental toxicology. The diversity in interactions described in the food web or food chain is not expressed in a term like the Shannon-Wiener index. However, in aquatic as well as soil ecological research, extensive, quantitative descriptions have been made of various ecosystems. These model descriptions, like the one for arable soil below, are partly based on the taxonomic background of species groups and partly on their functional role in the food web, expressed as their way of feeding (see for instance the phytophagous nematodes feeding from plants, the fungivorous nematodes eating fungi and the predaceous nematodes eating other nematodes).

The scheme in Figure 1 shows a very general soils food web and the different trophic levels. Much more detailed soil food web descriptions also are available, that do not only link the different trophic groups but also describe the energy flows within the system and through these flows the intensity and thus strength of the interactions that together determine the stability of the system (see e.g. de Ruiter et al., 1998).

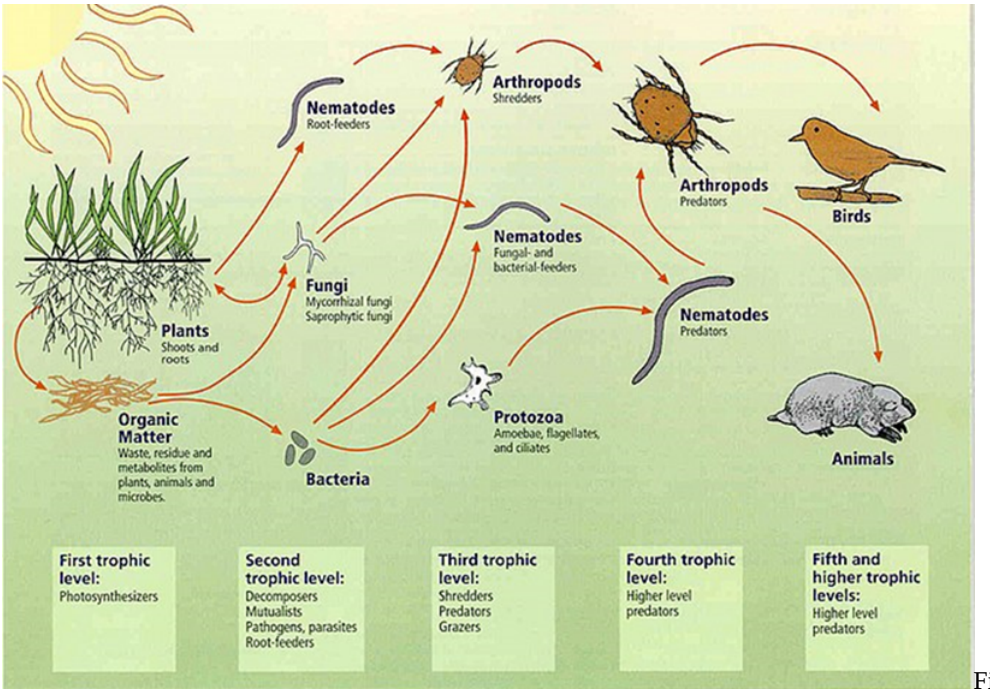


Figure 1: Simplified soil food web showing the relationships between different organisms living in and on the soils and their trophic relationships. See text for further explanation. Source: https://en.Wikipedia.org/wiki/Soil_food_web#/media/File:Soil_food_webUSDA.jpg

This food web shown in Figure 1 illustrates that biodiversity not only has a structural side: the various types of species, but also a functional one: which species are involved in the execution of which process. At the species level this functional aspect has been further elaborated in specific feeding leagues. At the ecosystem level this functional aspect has clearly been recognized in the last decades and resulted in the development of the concept of ecosystem services (see Section on Ecosystem services). However, these do not trace back to the individual species level and as such not to the functional aspect of biodiversity. Another development to be mentioned is that of trait-based approaches, which attempt to group species according to certain traits that are linked not only to exposure and sensitivity but also to their functioning. With that the trait-based approach may enable linking structural and functional biodiversity (see Section on Trait-based approaches).

Functional redundancy

When effects of contaminants on species are compared to effects on processes, the species effects are mostly more distinct than the process effects. In other words: effects on structural diversity will be seen already at lower concentrations, and probably also sooner, than effects on functional diversity. This can be explained by the fact that processes are executed by more than one species. When with increasing chemical exposure levels the most sensitive species disappear, their role is taken over by less sensitive species. This reasoning has been generalized in the concept of "*functional redundancy*", which postulates that not all species that can perform a specific process are always active, and thus necessary, in a specific situation. Consequently they are superfluous or redundant. When a sensitive species that can perform a similar function disappears, a redundant species may take over, so the function is still covered. It has to be realized, however, that in case this is relevant in situation A, that does not mean it is also relevant for situation B with different environmental conditions and another species composition. Another consequence of functional redundancy is that when functional biodiversity is affected, there is (severe) damage to structural biodiversity: most likely several important species will have gone extinct or are strongly inhibited.

Examples of the relation between structure and functioning

Redundant species are often less efficient in performing a certain function. Tyler (1984) observed in a gradient of Cu-contamination by a copper manufacturing plant in Gusum Sweden that specific enzyme functions as well as a general processes like mineralisation decreased faster than the total fungal biomass. (Figure 2b). The explanation was provided by experimental research by Ruhling et al. (1984) who selected a number of these micro-fungi in the field and tested them for their sensitivity for copper.

The various species showed different concentration-effect relationships but all going to zero, except for two species which increased in abundance at the higher concentration so that the total biomass stayed more or less the same (Figure 2A).

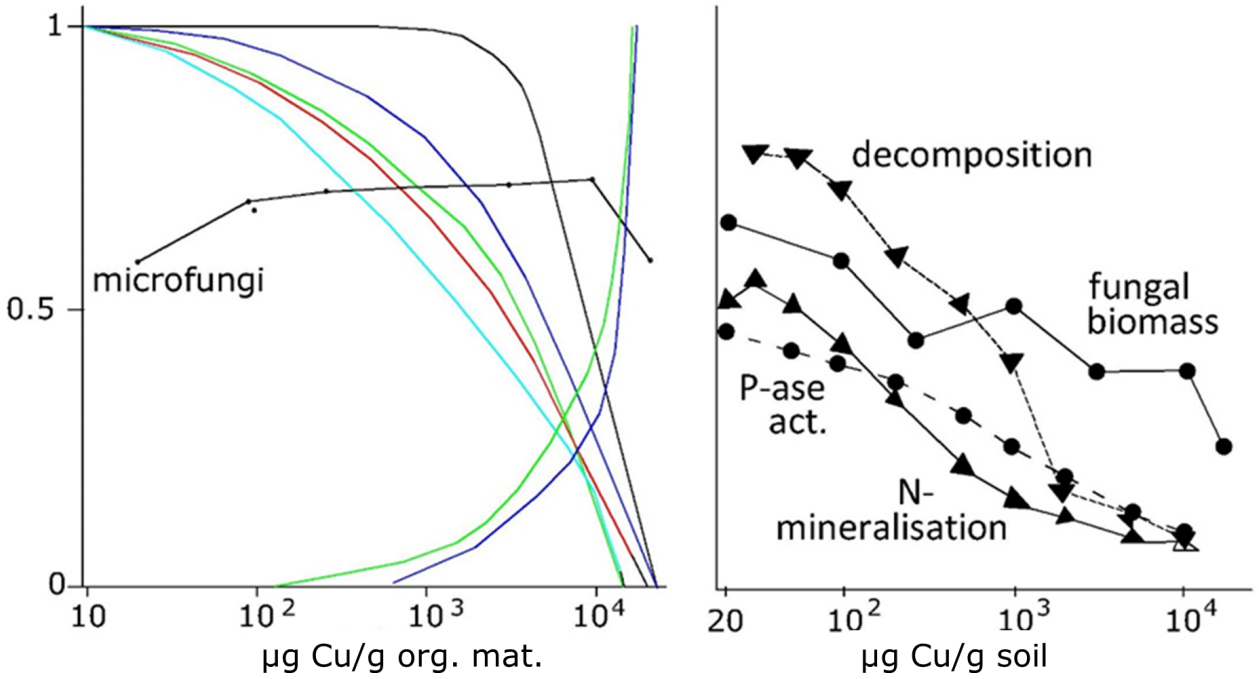


Figure 2. Left: Experimental dose responses of various microfungi species to increased levels of copper in the organic matter used as substrate and the total response for all tested species (Rühling et al., 1984). Right: Reduction of various breakdown processes and fungal biomass in the field with increasing copper levels in the soil (Tyler, 1984). Drawn by Wilma IJzerman.

Another example of the importance of a combined approach to structural and functional diversity are the different ecological types of earthworms. According to their behaviour and role they are distinguished in the anecics (deep burrowing earthworms moving up and down from deeper soil layers to the soil surface and consuming leaf litter), the endogeics (active in the deeper laying mineral and humus soil layers and consuming fragmented litter material and humus), and the epigeics (active in the upper soil litter layer and consuming leaf litter). Adverse effects of contamination on the anecics will result in accumulation of litter at the soil surface, in reduced litter fragmentation by the epigeics and reduced humus-forming by the endogeics. In various studies it is shown that these earthworms have different sensitivities for different types of pesticides. However, so far the ranking of more or less sensitive species is different for different groups of pesticides. So, there is no general relation between the function of a species e.g. surface active earthworms (epigeics) and their exposure to and sensitivity for pesticides. Nevertheless, pesticide effects on anecics generally lead to reduced litter removal, effects on endogeics result in slower fragmentation, reduced humification etc., and an effect on earthworm communities in general may hamper soil aeration and lead to soil compaction.

Another example of the impact of contaminants on functional diversity is from microbiological research on the impact of heavy metals by Doelman et al. (1994). They isolated fungi, bacteria and actinomycetes from various heavy metal contaminated and clean areas, tested these species for their sensitivity to zinc and cadmium, and divided them accordingly in a sensitive and resistant group. As a next step they measured to what extent both groups were able to degrade and mineralize a series of organic compounds. Figure 3 shows that the sensitive group is much more effective in degrading a variety of organic compounds, whereas the heavy metal resistant microbes are far less effective. This would indicate that although functional redundancy may alleviate some of the effects that contaminants have on ecosystem functioning, the overall performance of the community generally decreases upon contaminant exposure.

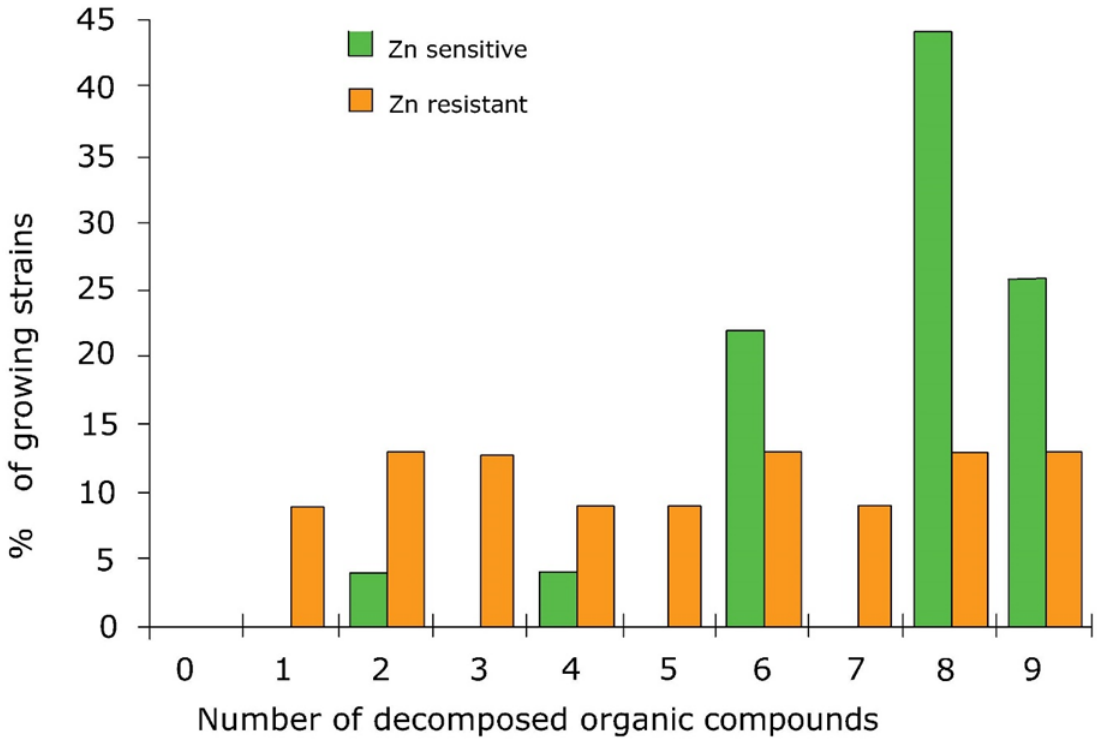


Figure 3.
Decomposing capacity (measured as growth) of Zn resistant bacteria and Zn sensitive bacteria for increasing numbers of organic compounds. Redrawn from Doelman et al. (1994) by Wilma Ijzerman.

The latter example also shows that genetic diversity, expressed as the numbers of sensitive and resistant species, plays a role in the functional stability and sustainability of microbial degradation processes in the soil.

In conclusion, ecosystem services are worth studying in relation to contamination (Faber et al., 2019), but also more specific in relation to functional diversity at the species level. A promising field of research in this framework would include microorganisms in relation to the variety of degradation processes they are involved in.

References

De Ruiter, J.C., Neutel, A-M., Moore, J.C. 1995. Energetics, patterns of interaction strenghts and stability in real ecosystems. Science 269, 1257-60.

Doelman, P., Jansen, E., Michels, M., Van Til, M. (1994). Effects of heavy metals in soil on microbial diversity and activity as shown by the sensitivity-resistance index, an ecologically relevant parameter Biology and Fertility of Soils 17, 177-184.

Faber, J.H., Marshall, S., Van den Brink, P.J., Maltby, L. (2019). Priorities and opportunities in the application of the ecosystem services concept in risk assessment for chemicals in the environment. Science of the Total Environment 651, 1067-1077.

Rühling, Å., Bååth, E., Nordergren, A., Söderström, B. (1984) Fungi in a metal-contaminated soil near the Gusum brass mill, Sweden. Ambio 13, 34-36.

Tyler, G. (1984) The impact of heavy metal pollution on forests: A case study of Gusum, Sweden. Ambio 13, 18-24.

5.8. Question 1

Describe the structural and functional diversity at the species and at the landscape level

---

5.8. Question 2

What is meant by redundancy?

5.8. Question 3

Does redundancy have an impact of the sensitivity of species (structural diversity) versus processes (functional diversity)?

This page titled 2.3.5.8: Structure versus function incl. ecosystem services is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **5.8: Structure versus function incl. ecosystem services** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.5.9: Landscape ecotoxicology

5.9. Landscape ecotoxicology

In preparation

This page titled 2.3.5.9: Landscape ecotoxicology is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **5.9: Landscape ecotoxicology** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

SECTION OVERVIEW

2.3.6: Risk Assessment and Regulation

2.3.6.1: Introduction - The Essence of Risk Assessment

2.3.6.2: Ecosystem services and protection goals

2.3.6.3: Predictive risk assessment approaches and tools

2.3.6.4: Diagnostic risk assessment approaches and tools

2.3.6.5: Regulatory Frameworks

2.3.6.6: Risk management and risk communication

2.3.6.7: Risk perception

This page titled 2.3.6: Risk Assessment and Regulation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

2.3.6.1: Introduction - The Essence of Risk Assessment

Author: Ad Ragas

Reviewer: Martien Janssen

Learning Objectives

After this module, you should be able to:

- explain the terms risk, hazard, risk assessment, risk management and solution-focused risk assessment ;
- explain the different steps of the risk assessment process, the relation between these steps and how the principle of tiering works;
- give an example of a risk indicator;
- indicate the most important advantages and disadvantages of the risk assessment paradigm.

Key words

Risk, hazard, tiering, problem definition, exposure assessment, effect assessment, risk characterization

Introduction

We assess risks on a daily basis, although we may not always be aware of it. For example, when we cross the street, we - often implicitly - assess the benefits of crossing and weigh these against the risks of getting hit by a vehicle. If the risks are considered too high, we may decide not to cross the street, or to walk a bit further and cross at a safer spot with traffic lights.

Risk assessment is common practice for a wide range of activities in society, for example for building bridges, protection against floods, insurance against theft and accidents, and the construction of a new industrial plant. The principle is always the same: we use the available knowledge to assess the probability of potential adverse effects of an activity as good as we can. And if these risks are considered too high, we consider options to reduce or avoid the risk.

Terminology

Risk assessment of chemicals aims to describe the risks resulting from the use of chemicals in our society. In chemical risk assessment, risk is commonly defined as "the probability of an adverse effect after exposure to a chemical". This is a very practical definition that provides natural scientists and engineers the opportunity to quantify risk using "objective" scientific methods, e.g. by quantifying exposure and the likelihood of adverse effects. However, it should be noted that this definition ignores more subjective aspects of risk, typically studied by social scientists, e.g. the perceptions of people and (dealing with) knowledge gaps. This subjective dimension can be important for risk management. For example, risk managers may decide to take action if a risk is perceived as high by a substantial part of the population, even if the associated health risks have been assessed as negligible by natural scientists and engineers.

Next to the term "risk", the term "hazard" is often used. The difference between both terms is subtle, but important. A hazard is defined as the inherent capacity of a chemical (or agent/activity) to cause adverse effects. The labelling of a substance as "carcinogenic" is an example of a hazard-based action. The inherent capacity of the substance to trigger cancer, as for example demonstrated in an *in vitro* assay or an experiment with rats or mice, can be sufficient reason to label a substance as "carcinogenic". Hazard is thus independent of the actual exposure level of a chemical, whereas risk is not.

Risk assessment is closely related to risk management, i.e. the process of dealing with risks in society. Decisions to accept or reduce risks belong to the risk management domain and involve consideration of the socio-economic implications of the risks as well as the risk management options. Whereas risk assessment is typically performed by natural scientists and engineers, often referred to as "risk assessors", risk management is performed by policy makers, often referred to as "risk managers".

Risk assessment and risk management are often depicted as sequential processes, where assessment precedes management. However, strict separation of both processes is not always possible and management decisions may be needed before risks are assessed. For example, risk assessment requires political agreement on what should be protected and at what level, which is a risk management issue (see Section on Protection Goals). Similarly, the identification, description and assessment of uncertainties in the

assessment is an activity that involves risk assessors as well as risk managers. Finally, it is often more efficient to define alternative management options before performing a risk assessment. This enables the assessment of the current situation and alternative management scenarios (i.e., potential solutions) in one round. The scenario with the maximum risk reduction that is also feasible in practice would then be the preferred management option. This mapping of solutions and concurrent assessment of the associated risks is also known as solution-focused risk assessment.

Risk assessment steps and tiering

Chemical risk assessment is typically organized in a limited number of steps, which may vary depending on the regulatory context. Here, we distinguish four steps (Figure1):

1. Problem definition (sometimes also called hazard identification), during which the scope of the assessment is defined;
2. Exposure assessment, during which the extent of exposure is quantified;
3. Effect assessment (sometimes also called hazard or dose-response assessment), during which the relationship between exposure and effects is established;
4. Risk characterization, during which the results of the exposure and effect assessments are combined into an estimate of risk and the uncertainty of this estimate is described.

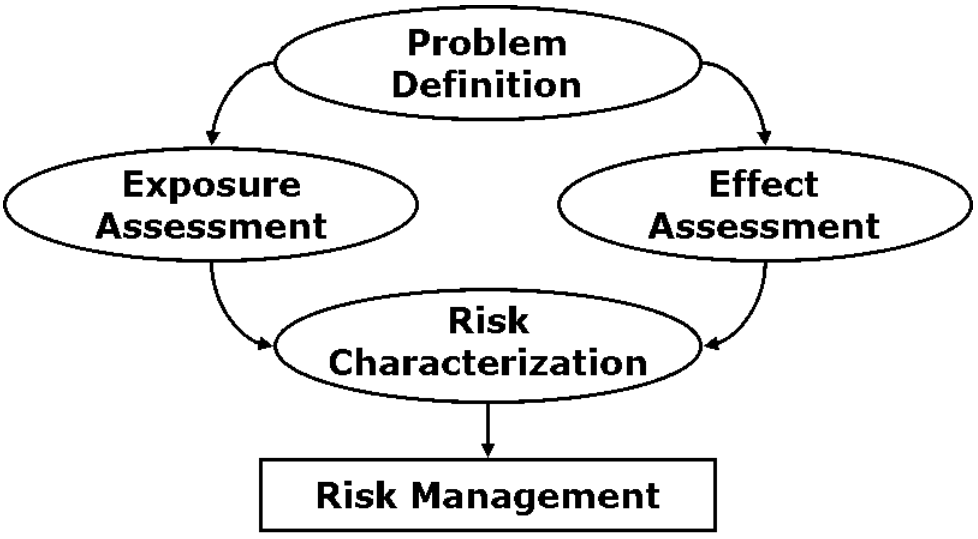


Figure 1. Risk assessment consists of four steps (problem definition, exposure assessment, effect assessment & risk characterization) and provides input for risk management.

The four risk assessment steps are explained in more detail below. The four steps are often repeated multiple times before a final conclusion on the acceptability of the risk is reached. This repetition is called tiering (Figure 2). It typically starts with a simple, conservative assessment and then, in subsequent tiers, more data are added to the assessment resulting in less conservative assumptions and risk estimates. Tiering is used to focus the available time and resources for assessing risks on those chemicals that potentially lead to unacceptable risks. Detailed data are gathered only for chemicals showing potential risk in the lower, more conservative tiers.

The order of the exposure and effect assessment steps has been a topic of debate among risk assessors and managers. Some argue that effect assessment should precede exposure assessment because effect information is independent of the exposure scenario and can be used to decide how exposure should be determined, e.g., information on toxicokinetics can be relevant to determine the exposure duration of interest. Others argue that exposure should precede effect assessment since assessing effects is expensive and unnecessary if exposure is negligible. The current consensus is that the preferred order should be determined on a case-by-case basis with parallel assessment of exposure and effects and exchange of information between the two steps as the preferred option.

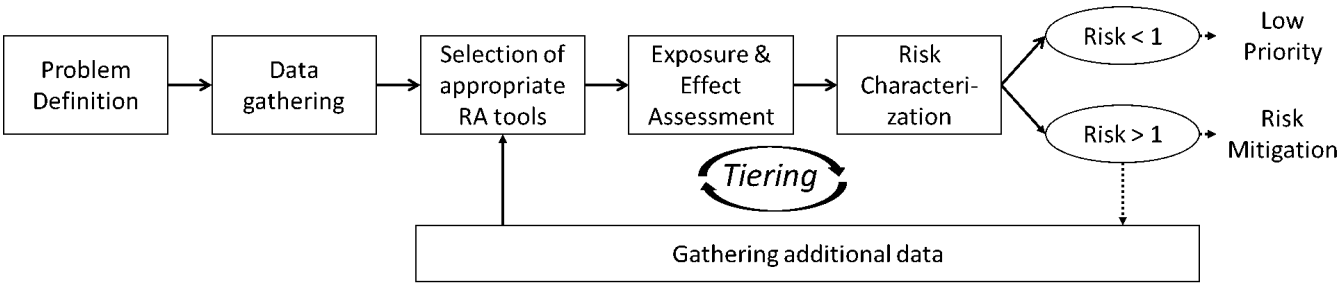


Figure 2: The principle of tiering in risk assessment. Initially risks are assessed using limited data and conservative assumptions and tools. When the predicted risk turns out unacceptable (Risk >1; see below), more data are gathered and less conservative tools are used.

Problem definition

The scope of the assessment is determined during the problem definition phase. Questions typically answered in the problem definition include:

- What is the nature of the problem and which chemical(s) is/are involved?
- What should be protected, e.g. the general population, specific sensitive target groups, aquatic ecosystems, terrestrial ecosystems or particular species, and at what level?
- What information is already available, e.g. from previous assessments?
- What are the available resources for the assessment?
- What is the assessment order and will tiering be applied?
- What exposure routes will be considered?
- What is the timeframe of the assessment, e.g. are acute or (sub)chronic exposures considered?
- What risk metric will be used to express the risk?
- How will uncertainties be addressed?

Problem definition is not a task for risk assessors only, but should preferably be performed in a collaborative effort between risk managers, risk assessors and stakeholders. The problem definition should try to capture the worries of stakeholders as good as possible. This is not always an easy task as these worries may be very broad and sometimes also poorly articulated. Risk assessors need a clearly demarcated problem and they can only assess those aspects for which assessment methods are available. The dialogue should make transparent which aspects of the stakeholder concerns will be assessed and which not. Being transparent about this can avoid disappointments later in the process, e.g. if aspects considered important by stakeholders were not accounted for because suitable risk assessment methods were lacking. For example, if stakeholders are worried about the acute and chronic impacts of pesticide exposure, but only the acute impacts will be addressed, this should be made clear at the beginning of the assessment.

The problem definition phase results in a risk assessment plan detailing how the risks will be assessed given the available resources and within the available timeframe.

Exposure assessment

An important aspect of exposure assessment is the determination of an exposure scenario. An exposure scenario describes the situation for which the exposure is being assessed. In some cases, this exposure situation may be evident, e.g. soil organisms living a contaminated site. However, especially when we want to assess potential risks of future substance applications, we have to come up a typical exposure scenario. Such scenarios are for example defined before a substance is allowed to be used as a food additive or before a new pesticide is allowed on the market. Exposure scenarios are often conservative, meaning that the resulting exposure estimate will be higher than the expected average exposure.

The exposure metric used to assess the risk depends on the protection target. For ecosystems, a medium concentration is often used such as the water concentration for aquatic systems, the sediment concentration for benthic systems and the soil concentration for terrestrial systems. These concentrations can either be measured or predicted using a fate model (see Section 3.8) and may or may not take into account bioavailability (see Section 3.6). For human risk assessment, the exposure metric depends on the exposure

route. An air concentration is often used to cover inhalation, the average daily intake from food and water to cover oral exposure, and uptake through skin for dermal exposure. Uptake through multiple routes can also be combined in a dose metric for internal exposure, such as Area Under the Curve (AUC) in blood (see Section 6.3.1). Exposure metrics for specific wildlife species (e.g. top predators) and farm animals are often similar as those for humans. Measuring and modelling route-specific exposures is generally more complex than quantifying a simple medium concentration, because it does not only require the quantification of the substance concentration in the contact medium (e.g., concentration in drinking water), but also quantification of the contact intensity (e.g., how much water is consumed per day). Especially oral exposure can be difficult to quantify because it covers a wide range of different contact media (e.g. food products) and intensities varying from organism to organism.

Effect assessment

The aim of the effect assessment is to estimate a reference exposure level, typically an exposure level which is expected to cause no or very limited adverse effects. There are many different types of reference levels in chemical risk assessment; each used in a different context. The most common reference level for ecological risk assessment is the Predicted No Effect Concentration (PNEC). This is the water, soil, sediment or air concentration at which no adverse effects at the ecosystem level are being expected. In human risk assessment, a myriad of different reference levels are being used, e.g. the Acceptable Daily Intake (ADI), the oral and inhalatory Reference Dose (RfD), the Derived No Effect Level (DNEL), the Point of Departure (PoD) and the Virtually Safe Dose (VSD). Each of these reference levels is used in a specific context, e.g. for addressing a specific exposure route (ADI is oral), regulatory domain (the DNEL is used in the EU for REACH, whereas the RfD is used in the USA), substance type (the VSD is typical for genotoxic carcinogens) or risk assessment method (the PoD is typical for the Margin of Safety approach).

What all reference levels have in common, is that they reflect a certain level of protection for a specific protection goal. In ecological risk assessment, the protection goal typically is the ecosystem, but it can also be a specific species or even an organism. In human risk assessment, the protection goal typically comprises all individuals of the human population. The definition of protection goals is a normative issue and it therefore is not a task of risk assessors, but of politicians. The protection levels defined by politicians typically involve a high level of abstraction, e.g. "the entire ecosystem and all individuals of the human population should be protected". Such abstract protection goals do not always match with the methods used to assess the risks. For example, if one assumes that one molecule of a genotoxic carcinogen can trigger a deathly tumour, 100% protection for all individuals of the human population is feasible only by banning all genotoxic carcinogens (reference level = 0). Likewise, the safe concentration for an ecosystem is infinitely small if one assumes that the sensitivity of the species in the system follows a lognormal distribution which asymptotically approaches the x-axis. Hence, the abstract protection goals have to be operationalized, i.e. defined in more practical terms and matching the methods used for assessing effects. This is often done in a dialogue between scientific experts and risk managers. An example is the "*one in a million lifetime risk estimated with a conservative dose response model*" which is used by many (inter)national organizations as a basis for setting reference levels for genotoxic carcinogens. Likewise, the concentration at which the no observed effect concentration (NOEC) for only 5% of the species is being exceeded is often used as a basis for deriving a PNEC.

Once a protection goal has been operationalized, it must be translated into a corresponding exposure level, i.e. the reference level. This is typically done using the outcomes of (eco)toxicity tests, i.e. tests with laboratory animals such as rats, mice and dogs for human reference levels and with primary consumers, invertebrates and vertebrates for ecological reference levels. Often, the toxicity data are plotted in a graph with the exposure level on the x-axis and the effect or response level on the y-axis. A mathematical function is then fitted to the data; the so-called dose-response relationship. This dose-response relationship is subsequently used to derive an exposure level that corresponds to a predefined effect or response level. Finally, this exposure level is extrapolated to the ultimate protection goal, accounting for phenomena such as differences in sensitivity between laboratory and field conditions, between tested species and the species to be protected, and the (often very large) variability in sensitivity in the human population or the ecosystem. This extrapolation is done by dividing the exposure level that corresponds to a predefined effect or response level by one or more assessment or safety factors. These assessment factors do not have a pure scientific basis in the sense that they account for physiological differences which have actually been proven to exist. These factors also account for uncertainties in the assessment and should make sure that the derived reference level is a conservative estimate. The determination of reference levels is an art in itself and is further explained in sections 6.3.1 for human risk assessment and 6.3.2 for ecological risk assessment.

Risk characterization

The aim of risk characterization is to come up with a risk estimate, including associated uncertainties. A comparison of the actual exposure level with the reference level provides an indication of the risk:

$$Risk\ Indicator = \frac{Exposure\ Level}{Reference\ Level}$$

If the reference level reflects the maximum safe exposure level, then the risk indicator should be below unity (1.0). A risk indicator higher than 1.0 indicates a potential risk. It is a "potential risk" because many conservative assumptions may have been made in the exposure and effect assessments. A risk indicator above 1.0 can thus lead to two different management actions: (1) if available resources (time, money) allow and the assessment was conservative, additional data may be gathered and a higher tier assessment may be performed, or (2) consideration of mitigation options to reduce the risk. Assessment of the uncertainties is very important in this phase, as it reveals how conservative the assessment was and how it can be improved by gathering additional data or applying more advanced risk assessment tools.

Risks can also be estimated using a margin-of-safety approach. In this approach, the reference level used has not yet been extrapolated from the tested species to the protection goal, e.g. by applying assessment factors for interspecies and interindividual differences in sensitivity. As such, the reference level is not a conservative estimate. In this case, the risk indicator reflects the "margin of safety" between actual exposure and the non-extrapolated reference level. Depending on the situation at hand, the margin-of-safety typically should be 100 or higher. The main difference between the traditional and the margin-of-safety approach in risk assessment is the timing for addressing the uncertainties in the effect assessment.

Reflection

Figure 3 illustrates the risk assessment paradigm using the DPSIR chain (Section 1.2). It illustrates how reference exposure levels are being derived from protection goals, i.e. the maximum level of impact that we consider acceptable. The actual exposure level is either measured or predicted using estimated emission levels and dispersion models. When measured exposure levels are used, this is called retrospective or diagnostic risk assessment: the environmental is already polluted and the assessor wants to know whether the risk is acceptable and which substances are contributing to it. When the environment is not yet polluted, predictive tools can be used. This is called prospective risk assessment: the assessor wants to know whether a projected activity will result in unacceptable risks. Even if the environment is already polluted, the risk assessor may still decide to prefer predicted over measured exposure levels, e.g. if measurements are too expensive. This is possible only if the pollution sources are well-characterized. Retrospective (diagnostic) and prospective risk assessments can differ substantially in terms of problem definitions and methods used, and are therefore discussed in separate sections in this online book.

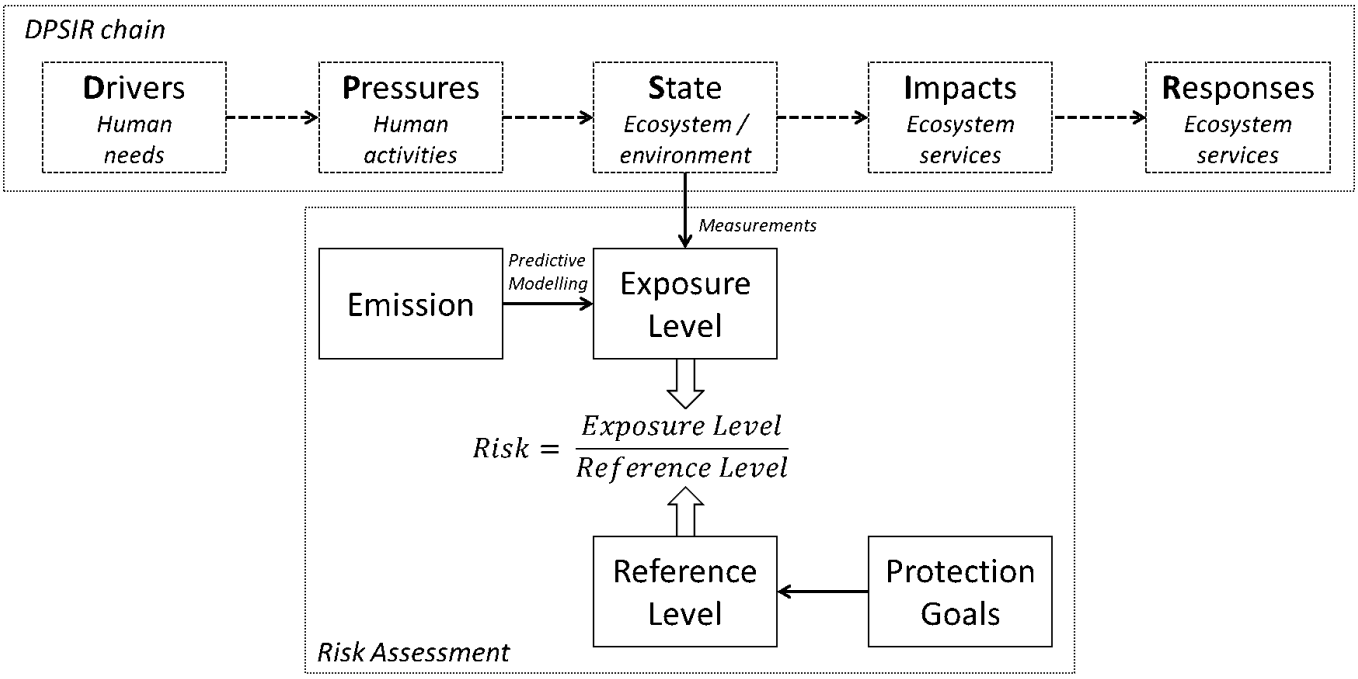


Figure 3: The risk assessment paradigm and the DPSIR chain.

Figure 3 can also be used to illustrate some important criticism on the current risk assessment paradigm, i.e. the comparison between the actual exposure level and a reference level. In current assessments, only one point of the dose-response relationship is being used to assess risk, i.e. the reference level. Critics argue that this is suboptimal and a waste of resources because the dose-response information is not used to assess the actual risk. A risk indicator with a value of 2.0 implies that the exposure is twice as high as the reference level but this does not give an indication of how many individuals or species are being affected or of the intensity of the effect. If the dose-response relationship would be used to determine the risk, this would result in a better-informed risk estimate.

A final critical remark that should be made, is the fact that risk assessment is often performed on a substance-by-substance basis. Dealing with mixtures of chemicals is difficult because each mixture has a unique composition in terms of compounds and concentration ratios between compounds. This makes it difficult to determine a reference level for mixtures. Mixture toxicology is slowly progressing and several methods are now available to address mixtures, i.e. whole mixture methods and compound-based approaches (Section 6.3.6). Another promising development are effect-based methods (Section 6.4.2). These methods do not assess risk based on chemical concentration, but on the toxicity measured in an environmental sample. In terms of DPSIR, these methods are assessing risks on the level of impacts rather than the level of state or pressures.

6.1. Question 1

Imagine the herbicide glyphosate would be banned based on its carcinogenic properties. Would this intervention be risk-based or hazard-based?

6.1. Question 2

Indicate whether the following activities should involve risk assessors, risk managers/politicians and/or stakeholders:

1. Determination of a safe dose level based on established protection goals;
2. Determination of protection goals;
3. Determination of intervention options;
4. Demarcation of the risk assessment problem;
5. Translation of abstract protection goals into operational goals.

6.1. Question 3

Indicate whether the following risk assessments are retrospective or prospective:

1. Determining the adverse impacts of a contaminated area on human health and the environment;
2. Quantifying the human health risk of current air pollution levels;
3. Determining whether the risks associated with a new pesticide are acceptable;
4. Predicting the risk of chemicals based on current emission levels.

6.1. Question 4

A risk assessment was performed for two different substances, i.e. A and B. The risk indicator value of substance A was 1.5 and that of substance B was 2.0. A risk manager proposes to first address substance B and subsequently substance A. Do you agree? Motivate your answer.

This page titled 2.3.6.1: Introduction - The Essence of Risk Assessment is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **6.1: Introduction - The Essence of Risk Assessment** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.6.2: Ecosystem services and protection goals

6.2. Ecosystem services and protection goals

In preparation

This page titled 2.3.6.2: Ecosystem services and protection goals is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **6.2: Ecosystem services and protection goals** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.6.3: Predictive risk assessment approaches and tools

6.3. Predictive risk assessment approaches and tools

6.3.1. Environmental realistic scenarios (PECs) – Human

under review

6.3.2. Environmental realistic scenarios (PECs) – Eco

Authors: Jos Boesten, Theo Brock

Reviewer: Ad Ragas, Andreu Rico

Learning objectives:

You should be able to:

- explain the role of exposure scenarios in environmental risk assessment (ERA)
- describe the need for, and basic principles of defining exposure assessment goal
- link exposure and effect assessments and can describe the role of environmental scenarios in future ERAs

Keywords: pesticides, exposure, scenarios, assessment goals, effects

Role of exposure scenarios in environmental risk assessment (ERA)

An exposure scenario describes the combination of circumstances needed to estimate exposure by means of models. For example, scenarios for modelling pesticides exposure can be defined as a combination of abiotic (e.g. properties and dimensions of the receiving environment and related soil, hydrological and climate characteristics) and agronomic (e.g. crops and related pesticide application) parameters that are thought to represent a realistic worst-case situation for the environmental context in which the exposure model is to be run. A scenario for exposure of aquatic organisms could be e.g. a ditch with a minimum water depth of 30 cm alongside a crop growing on a clay soil with annual applications of pesticide using a 20-year time series of weather data and including pesticide exposure via spray drift deposition and leaching from drainpipes. Such a scenario would require modelling of spray drift, leaching from drainpipes and exposure in surface water, ending up in a 20-year time series of the exposure concentration. In this chapter, we explain the use of exposure scenarios in prospective ERA by giving examples for the regulatory assessment of pesticides in particular.

Need for defining exposure assessment goals

Between about 1995 and 2001 groundwater and surface water scenarios were developed for EU pesticide registration; also referred to as the FOCUS scenarios. The European Commission indicated that these should represent 'realistic worst-cases', a political concept which leaves considerable room for scientific interpretation. Risk assessors and managers agreed that the intention was to generate 90th percentile exposure concentrations. The concept of a 90th percentile exposure concentration assumes a statistical population of concentrations and 90% of these concentrations are lower than this 90th percentile (and thus 10% are higher). This 90th percentile approach has since then been followed for most environmental exposure assessments for pesticides at EU level.

The selection of the FOCUS groundwater and surface water scenarios involved a considerable amount of expert judgement because this selection could not yet be based on well-defined GIS procedures and databases on properties of the receiving environment. The EFSA exposure assessment for soil organisms was the first environmental exposure assessment that could be based on a well-defined GIS procedure, using EU maps of parameters like soil organic matter, density of crops and weather. During the development of this exposure assessment, it became clear that the concept of a 90th percentile exposure concentration is too vague: it is essential to define also the statistical population of concentrations from which this 90th percentile is taken. Based on this insight, the EFSA Panel on Plant Protection Products and their Residues (PPR) developed the concept of the exposure assessment goals, which has become the standard within EFSA for developing regulatory exposure scenarios for pesticides.

Procedure for defining exposure assessment goals

Figure 1 shows how an exposure assessment goal for the risk assessment of aquatic organisms can be defined following this EFSA procedure. The left part specifies the temporal dimensions and the right part the spatial dimensions. In box E1, the Ecotoxicologically Relevant type of Concentration (ERC) is defined, e.g. the freely dissolved pesticide concentration in water for pelagic organisms. In box E2, the temporal dimension of this concentration is defined, e.g. annual peak or time-weighted average concentration for a pre-defined period. Based on these elements, the multi-year temporal population of concentrations can be generated for one single water body (E5) which would consist of e.g. 20 peak concentrations in case of a time series of 20 years. The spatial part requires definition of the type of water body (e.g. ditch, stream or pond; box E3) and the spatial dimension of this body (e.g. having a minimum water depth of 30 cm; box E4). Based on these, the spatial population of water bodies can be defined (box E6), e.g. all ditches with a minimum water depth of 30 cm alongside fields treated with the pesticide. Finally, then in box E7 the percentile combination to be taken from the spatial-temporal population of concentrations is defined. Specification of the exposure assessment goals does not only involve scientific information, but also political choices because this specification influences the strictness of the exposure assessment. For instance, in case of exposure via spray drift a minimum water depth of 30 cm in box E4 leads to about a three times lower peak concentration in the water than a minimum water depth of 10 cm.

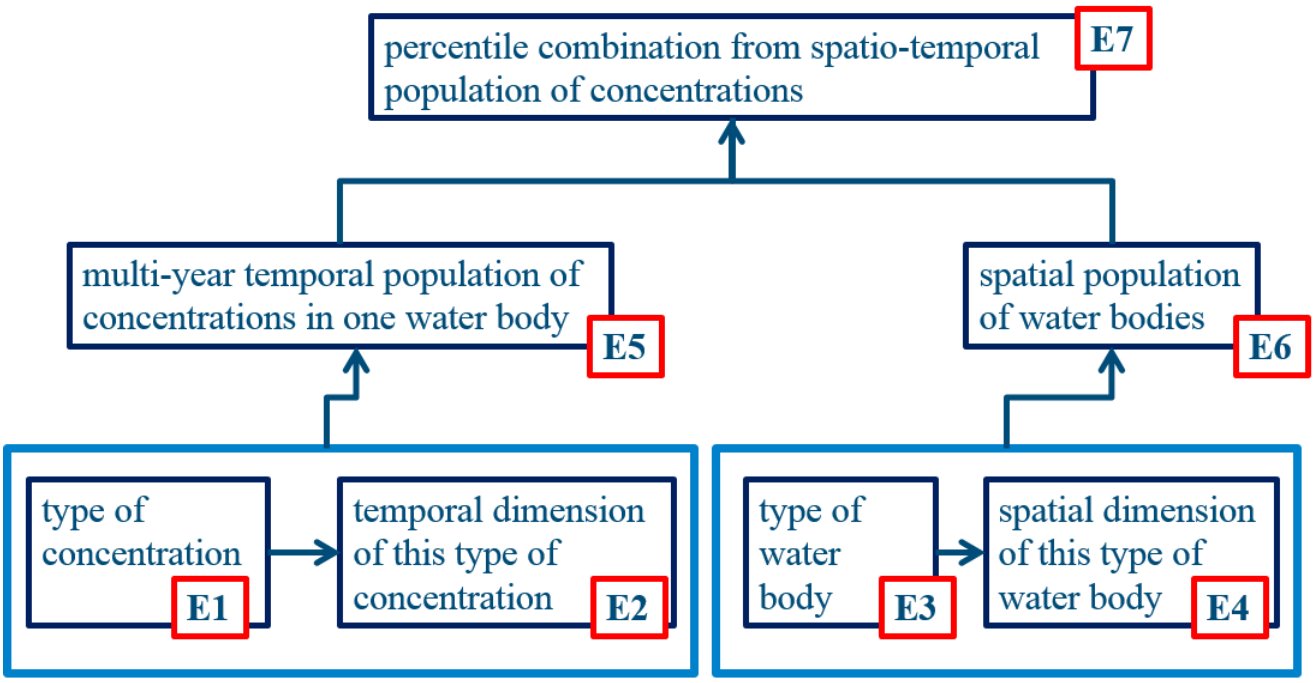


Figure 1. Scheme of the seven elements of the exposure assessment goal for aquatic organisms.

The schematic approach of Figure 1 can easily be adapted to other exposure assessment goals.

Interaction between exposure and effect assessment for organisms

Nearly all the environmental protection goals for pesticides involve assessment of risk for organisms; only groundwater and drinking water from surface water are based on a concentration of 0.1 µg/L which is not related to possible ecotoxicological effects. The risk assessment for organisms is a combination of an exposure assessment and an effect assessment as is illustrated by Figure 2.

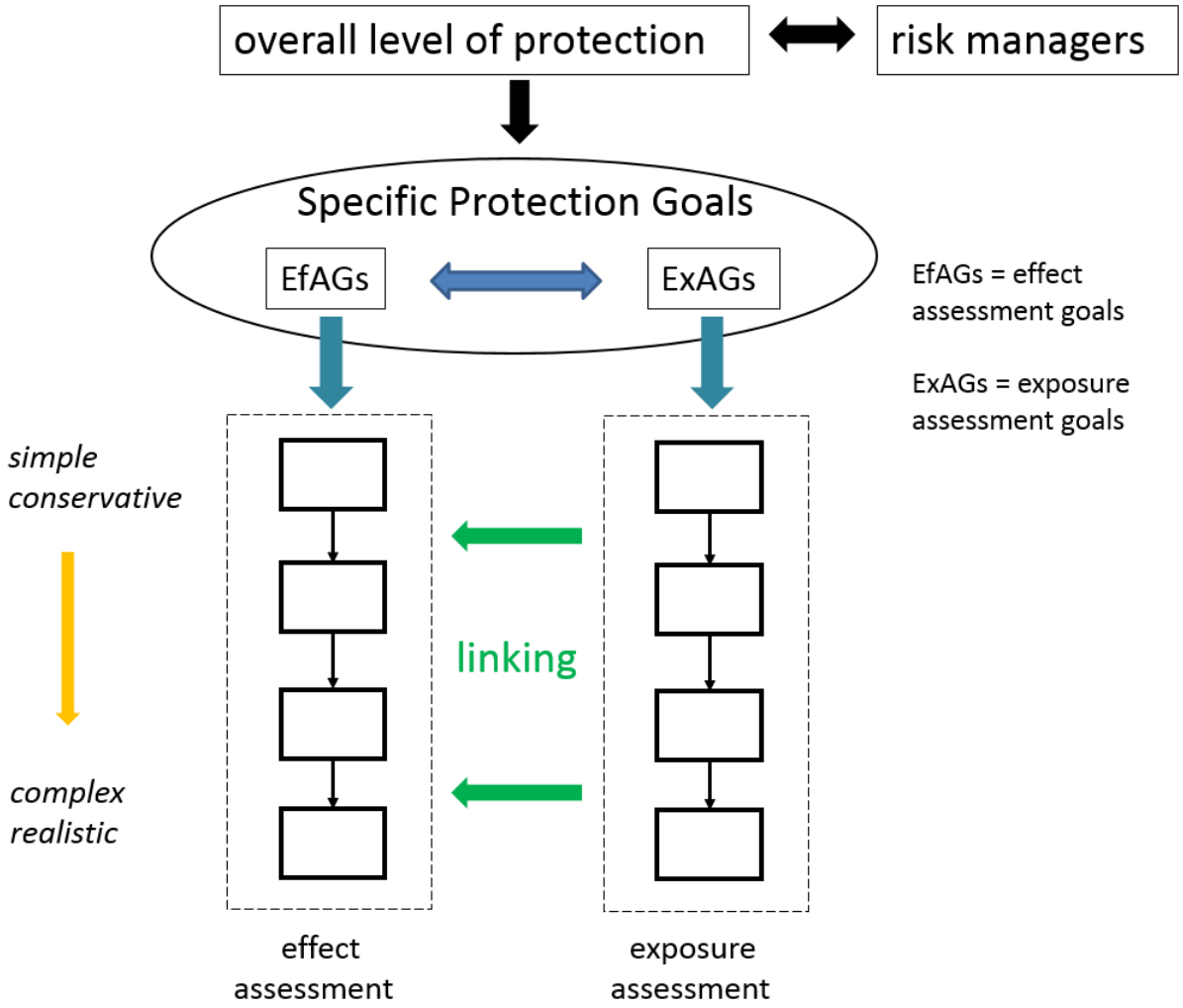


Figure 2. Overview of the risk assessment of organisms based on parallel tiered effect and exposure assessments.

Both the effect and the exposure assessment are tiered approaches with simple and conservative first tiers and less simple and more realistic higher tiers. A lower exposure tier may consist of a simple conservative scenario whereas a higher exposure tier may e.g. be based on a scenario selected using sophisticated spatial modelling. The top part of the scheme shows the link to the risk managers which are responsible for the overall level of protection. This overall level of protection is linked to the so-called Specific Protection Goals which will be explained in Section 6.5.3 and form the basis for the definition of the effect and exposure assessment goals. So the exposure assessment goals and resulting exposure scenarios should be consistent with the Specific Protection Goals (e.g. algae and fish may require different scenarios). When linking the two assessments, it has to be ensured that the type of concentration delivered by the exposure assessment is consistent with that required by the effect assessment (e.g. do not use time-weighted average concentrations in acute effect assessment). Figure 2 shows that in the assessment procedure information flows always from the exposure assessment to the effect assessment because the risk assessment conclusion is based on the effect assessment.

A relatively new development is to assess exposure and effects at the landscape level. This typically is a combination of higher-tier effect and exposure assessments. In such an approach, first the dynamics in exposure is assessed for the full landscape, and then combined with the dynamics of effects, for example based on spatially-explicit population models for species typical for that landscape. Such an approach makes a separate definition of the exposure and effect scenario redundant because this approach aims to deliver the exposure and effect assessment in an integrated way in space and time. Such an integrated approach requires the definition of "environmental scenarios". Environmental scenarios integrate both the parameters needed to define the exposure (exposure scenario) and those needed to calculate direct and indirect effects and recovery (ecological scenario) (see Figure 3).

However, it will probably take at least a decade before landscape-level approaches, including agreed-upon environmental scenarios, will be implemented for regulatory use in prospective ERA.

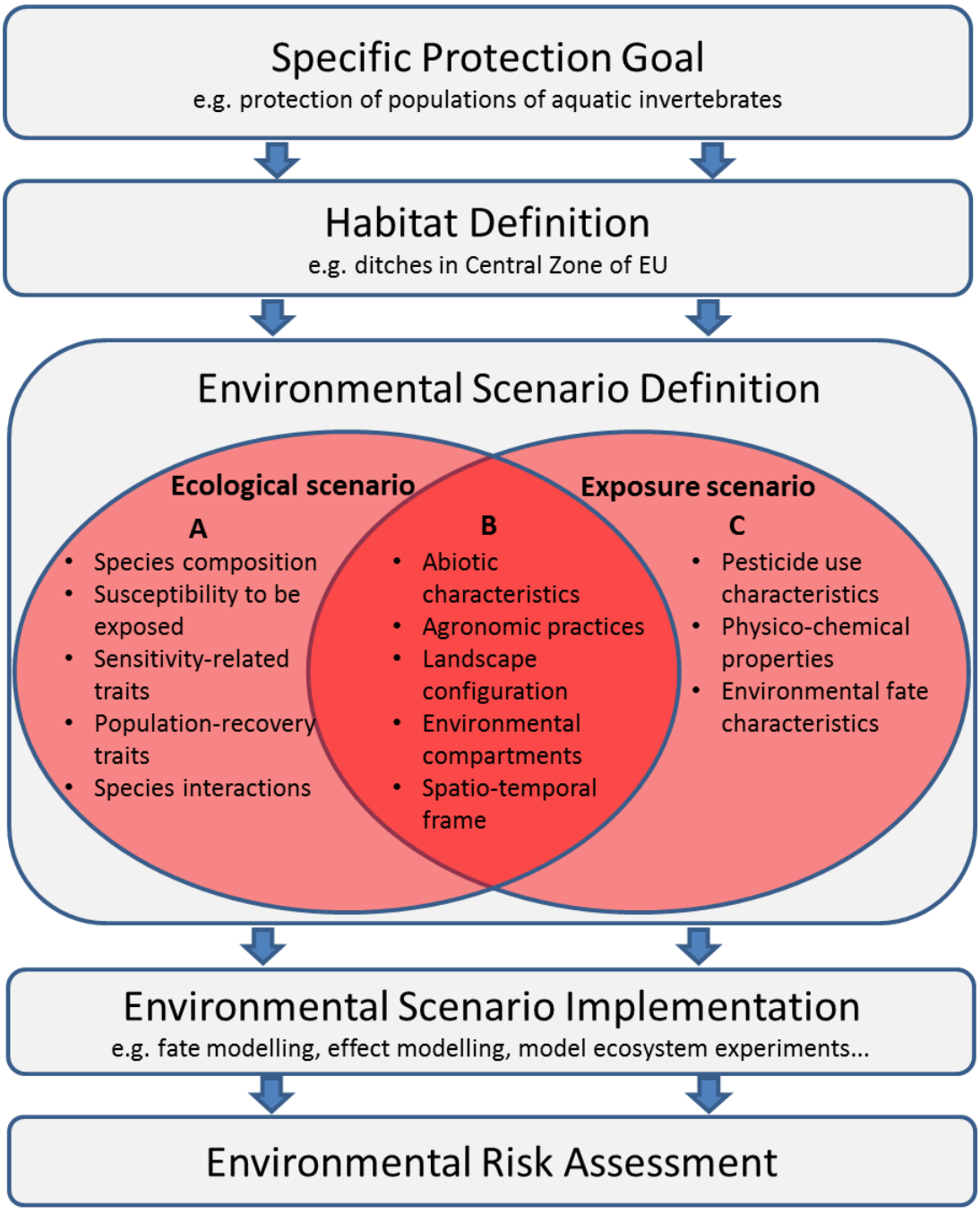


Figure 3. Conceptual framework of the role of an environmental scenario in prospective ERA (adapted after Rico et al. 2016).

References

Boesten, J.J.T.I. (2017). Conceptual considerations on exposure assessment goals for aquatic pesticide risks at EU level. Pest Management Science 74, 264-274.

Brock, T.C.M., Alix, A., Brown, C.D., et al. (2010). Linking aquatic exposure and effects: risk assessment of pesticides. SETAC Press & CRC Press, Taylor & Francis Group, Boca Raton, FL, 398 pp.

Rico, A., Van den Brink, P.J., Gylstra, R., Focks, A., Brock, T.C.M. (2016). Developing ecological scenarios for the prospective aquatic risk assessment of pesticides. Integrated Environmental Assessment and Management 12, 510-521.

6.3.2. Question 1

Why is a detailed specification of exposure assessment goals needed ?

6.3.2. Question 2

Why does specification of the exposure assessment goals include political choices ?

6.3.2. Question 3

Why does the risk assessment of organisms consist of two parallel tiered schemes for effects and exposure ?

6.3.3. Setting reference levels for human health protection

in preparation

6.3.4. Setting safe standards

Authors: Els Smit, Eric Verbruggen

Reviewers: Alexandra Kroll, Inge Werner

Learning objectives

You should be able to:

- explain what a reference level for ecosystem protection is;
- explain the basic concepts underlying the assessment factor approach for deriving PNECs;
- explain why secondary poisoning needs specific consideration when deriving a PNEC using the assessment factor approach.

Key words: PNEC, quality standards, extrapolation, assessment factor

Introduction

The key question in environmental risk assessment is whether environmental exposure to chemicals leads to unacceptable risks for human and ecosystem health. This is done by comparing the measured or predicted concentrations in water, soil, sediment, or air, with a reference level. Reference levels represent a dose (intake rate) or concentration in water, soil, sediment or air below which unacceptable effects are not expected. The definition of 'no unacceptable effects' may differ between regulatory frameworks, depending on the protection goal. The focus of this section is the derivation of reference levels for aquatic ecosystems as well as for predators feeding on exposed aquatic species (secondary poisoning), but the derivation of reference values for other environmental compartments follows the same principles.

Terminology and concepts

Various technical terms are in use as reference values, e.g. the Predicted No Effect Concentration (PNEC) for ecosystems or the Acceptable Daily Intake (ADI) for humans (Section on Human toxicology). The term "reference level" is a broad and generic term, which can be used independently of the regulatory context or protection goal. In contrast, the term "quality standard" is associated with some kind of legal status, e.g., inclusion in environmental legislation like the Water Framework Directive (WFD). Other terms exist, such as the terms 'guideline value' or 'screening level' which are used in different countries to indicate triggers for further action. While the scientific basis of these reference values may be similar, their implementation and the consequences of exceedance are not. It is therefore very important to clearly define the context of the derivation and the terminology used when deriving and publishing reference levels.

PNEC

A frequently used reference level for ecosystem protection is the Predicted No Effect Concentration (PNEC). The PNEC is the concentration below which adverse effects on the ecosystem are not expected to occur. PNECs are derived per compartment and apply to the organisms that are directly exposed. In addition, for chemicals that accumulate in prey, PNECs for secondary poisoning of predatory birds and mammals are derived. The PNEC for direct ecotoxicity is usually based on results from single species laboratory toxicity tests. In some case, data from field studies or mesocosms may be included.

A basic PNEC derivation for the aquatic compartment is based on data from single species tests with algae, water fleas and fish. Effects on the level of a complex ecosystem are not fully represented by effects on isolated individuals or populations in a laboratory set-up. However, data from laboratory tests can be used to extrapolate to the ecosystem level if it is assumed that protection of ecosystem structure ensures protection of ecosystem functioning, and that effects on ecosystem structure can be predicted from species sensitivity.

Accounting for Extrapolation Uncertainty: Assessment Factor (AF) Approach

To account for the uncertainty in the extrapolation from single species laboratory tests to effects on real life ecosystems, the lowest available test result is divided by an assessment factor (AF). In establishing the size of the AF, a number of uncertainties must be addressed to extrapolate from single-species laboratory data to a multi-species ecosystem under field conditions. These uncertainties relate to intra- and inter-laboratory variation in toxicity data, variation within and between species (biological variance), test duration and differences between the controlled laboratory set-up and the variable field situation. The value of the AF depends on the number of studies, the diversity of species for which data are available, the type and duration of the experiments, and the purpose of the reference level. Different AFs are needed for reference levels for e.g. intermittent release, short-term concentration peaks or long-term (chronic) exposure. In particular, reference levels for intermittent release and short-term exposure may be derived on the basis of acute studies, but short-term tests are less predictive for a reference level for long-term exposure and larger AFs are needed to cover this. Table 1 shows the generic AF scheme that is used to derive PNECs for long-term exposure of freshwater organisms in the context of European regulatory framework for industrial chemicals (REACH; see Section on REACH environment). This scheme is also applied for the authorisation of biocidal products, pharmaceuticals and for derivation of long-term water quality standards for freshwater under the EU Water Framework Directive. Further details on the application of this scheme, e.g., how to compare acute and chronic data and how to deal with irregular datasets, are presented in guidance documents (see suggested reading: EC, 2018; ECHA, 2008). Similar schemes exist for marine waters, sediment, and soil. However, for the latter two environmental compartments often too little experimental information is available and risk limits have to be calculated by extrapolation from aquatic data using the Equilibrium Partitioning concept. The derivation of Regulatory Acceptable Concentrations (RAC) for plant protection products (PPPs) is also based on the extrapolation of laboratory data, but follows a different approach focussing on generating data for specific taxonomic groups, taking account of the mode of action of the PPP (see suggested reading: EFSA, 2013).

Table 1. Basic assessment factor scheme used for the derivation of PNECs for freshwater ecosystems used in several European regulatory frameworks. Consult the original guidance documents for full schemes and additional information (see suggested reading: EC, 2018; ECHA, 2008).

Available data	**Assessment factor**
At least one short-term L(E)C50 from each of three trophic levels (fish, invertebrates (preferred *Daphnia*) and algae)	1000
One long-term EC10 or NOEC (either fish or *Daphnia*)	100
Two long-term results (e.g. EC10 or NOECs) from species representing two trophic levels (fish and/or Daphnia and/or algae)	50
Long-term results (e.g. EC10 or NOECs) from at least three species (normally fish, *Daphnia* and algae) representing three trophic levels	10

Application of Species Sensitivity Distribution (SSD) and Other Additional Data

The AF approach was developed to account for the uncertainty arising from extrapolation from (potentially limited) experimental datasets. If enough data are available for other species than algae, daphnids and fish, statistical methods can be applied to derive a PNEC. Within the concept of species sensitivity distribution (SSD), the distribution of the sensitivity of the tested species is used to estimate the concentration at which 5% of all species in the ecosystem is affected (HC5; see section on SSDs). When used for

regulatory purposes in European regulatory frameworks, the dataset should meet certain requirements regarding the number of data points and the representation of taxa in the dataset, and an AF is applied to the HC5 to cover the remaining uncertainty from the extrapolation from lab to field.

Where available, results from semi-field experiments (mesocosms, see section on Community ecotoxicology) can also be used, either on its own or to underpin the PNEC derived from the AF or SSD approach. SSDs and mesocosm-studies are also used in the context of authorisation of PPP.

Reference levels for secondary poisoning

Substances might be toxic to wildlife because of bioaccumulation in prey or a high intrinsic toxicity to birds and mammals. If this is the case, a reference level for secondary poisoning is derived for a simple food chain: water è fish or mussel è predatory bird or mammal. The toxicity data from bird or mammal tests are transformed into safe concentrations in prey. This can be done by simply recalculating concentrations in laboratory feed into concentrations in fish using default conversion factors (see e.g., ECHA, 2008). For the derivation of water quality standards under the WFD, a more sophisticated method was introduced that uses knowledge on the energy demand of predators and energy content in their food to convert laboratory data to a field situation. Also, the inclusion of other, more complex and sometimes longer food chains is possible, for which field bioaccumulation factors are used rather than laboratory derived values.

Suggested additional reading

EC (2018). Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Guidance Document No. 27. Technical Guidance For Deriving Environmental Quality Standards. Updated version 2018. Brussels, Belgium. European Commission. circabc.europa.eu/ui/group/9...7a2a6b/details

ECHA (2008). Guidance on information requirements and chemical safety assessment Chapter R.10: Characterisation of dose [concentration]-response for environment. Helsinki, Finland. European Chemicals Agency. May 2008. https://echa.europa.eu/documents/10162/13632/information_requirements_r10_en.pdf/bb902be7-a503-4ab7-9036-d866b8ddce69

EFSA (2013). Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. EFSA Journal 2013; 11(7): 3290 efsa.onlinelibrary.wiley.com...efsa.2013.3290

Traas, T.P., Van Leeuwen, C. (2007). Ecotoxicological effects. In: Van Leeuwen, C., Vermeire, T.C. (Eds.). Risk Assessment of Chemicals: an Introduction, Chapter 7. Springer.

6.3.4. Question 1

What is a PNEC?

6.3.4. Question 2

How is a basic PNEC commonly derived in Europe?

6.3.4. Question 3

Why are assessment factors applied?

6.3.4. Question 4

Which aspects are covered by the assessment factor?

6.3.4. Question 5

Within the EU REACH/WFD regulatory framework, which assessment factor may be applied to derive a PNEC for freshwater if you have one LC_{50} value for *Oncorhynchus mykiss*, one EC_{50} value for *Daphnia magna*, one EC_{10} value for *Oncorhynchus mykiss*, and one NOEC value for *Pseudokirchneriella subcapitata*?

6.3.5. Species Sensitivity Distributions (SSDs)

Authors: Leo Posthuma, Dick de Zwart

Reviewers: Ad Ragas, Keith Solomon

Learning objectives:

You should be able to:

- explain that differences exist in the reaction of species to exposure to a chemicals;
- explain that these differences can be described by a statistical distribution;
- derive a Species Sensitivity Distribution (SSD) for sensitivity data;
- derive benchmark concentration from an SSD;
- derive a predicted impact from an SSD.

Keywords: Species Sensitivity Distribution (SSD), benchmark concentration, Potentially Affected Fraction of species (PAF)

Introduction

The relationship between dose or concentration (X) and response (Y) is key in risk assessment of chemicals (see section on Concentration-response relationships). Such relationships are often determined in laboratory toxicity tests; a selected species is exposed under controlled conditions to a series of increasing concentrations to determine endpoints such as the No Observed Effect Concentration (NOEC), the EC50 (the Effect Concentration causing 50% effect on a studied endpoint such as growth or reproduction), or the LC50 (the Effect Concentration causing 50% lethal effects). For ecological risk assessment, multiple species are typically tested to characterise the (variation in) sensitivities across species or taxonomic groups within the ecosystem. In the mid-1980s it had been observed that-like many natural phenomena-a set of ecotoxicity endpoint data, representing effect concentrations for various species, follows a bell-shaped statistical distribution. The cumulative distribution of these data is a sigmoid (S-shaped) curve. It was recognized, that this distribution had particular utility for assessing, managing and protecting environmental quality regarding chemicals. The bell-shaped distribution was thereupon named a Species Sensitivity Distribution (SSD). Since then, the use of SSD models has grown steadily. Currently, the model is used for various purposes, providing important information for decision-making.

Below, the dual utility of SSD models for environmental protection, assessment and management are shown first. Thereupon, the derivation and use of SSD models are elaborated in a stepwise sequence.

The dual utility of SSD models

A species sensitivity distribution (SSD) is a distribution describing the variance in sensitivity of multiple species exposed to a hazardous compound. The statistical distribution is often plotted using a log-scaled concentration axis (X), and a cumulative probability axis (Y, varying from 0 - 1; Figure 1).

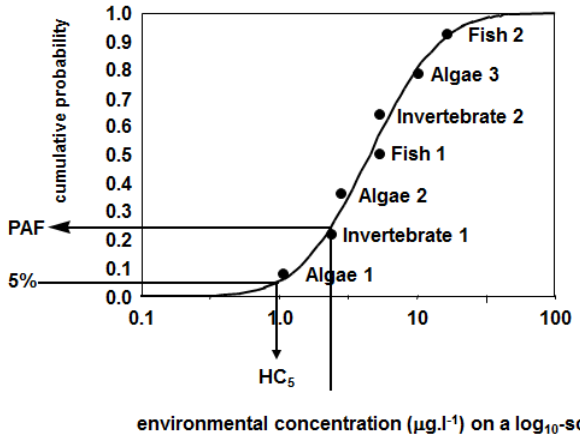


Figure 1. An species-sensitivity distribution (SSD) model, its data, and its dual use (from YàX, and from XàY). Dots represent the ecotoxicity endpoints (e.g., NOECs, EC50s, etc.) of different species.

Figure 1 shows that different species (here the dots represent 3 test data for algal species, 2 data for invertebrate species and 2 data fish species) have different sensitivities to the studied chemical. First, the ecotoxicity data are collected, and $\log_{10}$-transformed. Second, the data set can be visually inspected by plotting the bell-shaped distribution of the log-transformed data; deviations of the

expected bell-shape can be visually identified in this step. They may originate from causes such as a low number of data points or be indicative for a selective mode of action of the toxicant, such as a high sensitivity of insects to insecticides. Third, common statistical software for deriving the two parameters of the log-normal model (the mean and the standard deviation of the ecotoxicity data) can be applied, or the SSD can be described with a dedicated software tool such as ET_X (see below), including a formal evaluation of the 'goodness of fit' of the model to the data. With the estimated parameters, the fitted model can be plotted, and this is often done in the intuitively attractive form of the S-shaped cumulative distribution. This curve then serves two purposes. First, the curve can be used to derive a so-called Hazardous Concentration on the X-axis: a benchmark concentration that can be used as regulatory criterion to protect the environment (YàX). That is, chemicals with different toxicities have different SSDs, with the more hazardous compounds plotted to the left of the less hazardous compounds. By selecting a protection level on the Y-axis-representing a certain fraction of species affected, e.g. 5%-one derives the compound-specific concentration standards. Second, one can derive the fraction of tested species probably affected at an ambient concentration (XàY), which can be measured or modelled. Both uses are popular in contemporary environmental protection, risk assessment, and management.

Step 1: Ecotoxicity data for the derivation of an SSD model

The SSD model for a chemical and an environmental compartment (e.g., surface water, soil or sediment) is derived based on pertinent ecotoxicity data. Those are typically extracted from scientific literature or ecotoxicity databases. Examples of such databases are the U.S. EPA's Ecotox database, the European REACH data sets and the EnviroTox database which contains quality-evaluated studies. The researcher selects the chemical and the compartment of interest, and subsequently extracts all test data for the appropriate endpoint (e.g., ECx-values). The set of test data is tabulated and ranked from most to least sensitive. Multiple data for the same species are assessed for quality and only the best data are used. If there is > 1 toxicity value for a species after the selection process, the geometric mean value is commonly derived and used. A species should only be represented once in the SSD. Data are often available for frequently tested species, representing different taxonomic and/or trophic levels. A well-known triplet of species frequently tested is "Algae, Daphnids and Fish", as this triplet is a requested minimum set for various regulations in the realm of chemical safety assessment (see section on Regulatory frameworks). For various compounds, the number of test data can be more than hundred, whilst for most compounds few data of acceptable quality may be available.

Step 2. The derivation and evaluation of an SSD model

Standard statistical software (a spreadsheet program) or a dedicated software model such as ETX can be used to derive an SSD from available data. Commonly, the fit of the model to the data set is checked to avoid misinterpretation. Misfit may be shown using common statistical testing (Goodness of Fit tests) or by visual inspection and ecological interpretation of the data points. That is, when a chemical specifically affects one group of species (e.g., insects having a high sensitivity for insecticides), the user may decide the derive an SSD model for specific groups of species. In doing so, the outcome will consist of two or more SSDs for a single compound (e.g., an SSD_{Insect} and an SSD_{Other} when the compound is an insecticide, whilst the SSD_{Other} might be split further if appropriate). These may show a better goodness of fit of the model to the data, but - more importantly - they reflect the use of key knowledge of mode of action and biology prior to 'blindly' applying the model fit procedure.

Step 3a. The SSD model used for environmental protection

The oldest use of the SSD model is the derivation of reference levels such as the PNEC (YàX). That is, given the policy goal to fully protect ecosystems against adverse effects of chemical exposures (see Section on Ecosystem services and protection goals), the protective use is as follows. First, the user defines which ecotoxicity data are used. In the context of environmental protection, these have often been NOECs or low-effect levels (ECx, with low x, such as EC_{10}) from chronic tests. This yields an SSD-NOEC or SSD-ECx. Then, the user selects a level of Y, that is: the maximum fraction of species for which the defined ecotoxicity endpoint (NOEC or ECx) may be exceeded, e.g., 0.05 (a fraction of 0.05 equals 5% of the species). Next, the user derives the Hazardous Concentration for 5% of the species (YàX). At the HC_5, 5% of the species are exposed to concentrations *greater* than their NOEC, but-which is the obverse-95% of the species are exposed to concentration *less* than their NOEC. It is often assumed that the structural and functional integrity of ecosystems is sufficiently protected at the HC_5 level if the SSD is based on NOECs. Therefore, many authorities use this level to derive regulatory PNECs (Predicted No Effect Concentration) or Environmental Quality Standards (EQS). The latter concepts are used as official reference levels in risk assessment, the first is the preferred abbreviation in the context of prospective chemical safety assessments, and the second is used in retrospective environmental quality assessment. Sometimes an extra assessment factor varying between 1 and 5 is applied to the HC5 to account for remaining

uncertainties. Using SSDs for a set of compounds yields a set of HC_5 values, which-in fact-represent a relative ranking of the chemicals by their potential to cause harm.

Step 3b.The SSD model used for environmental quality assessment

The SSD model also can be used to explore how much damage is caused by environmental pollution. In this case, a predicted or measured ambient concentration is used to derive a Potentially Affected Fraction of species (PAF). The fraction ranges from 0-1 but, in practice, it is often expressed as a percentage (e.g., "24% of the species is likely affected"). According to this approach, users often have monitored or modelled exposure data from various water bodies, or soil or sediment samples, so that they can evaluate whether any of the studied samples contain a concentration higher than the regulatory reference level (previous section) and, if so how many species are affected. Evidently, the user must clearly express what type of damage is quantified, as damage estimates based on an SSD_{NOEC} or an SSD_{EC50} quantify the fractions of species affected beyond the no effect level and at the 50% effect level, respectively. This use of SSDs for a set of environmental samples yields a set of PAF values, which, in fact, represent a relative ranking of the pollution levels at the different sites in their potential to cause harm.

Practical uses of using SSD model outcomes

SSD model outcomes are used in various regulatory and practical contexts.

1. The oldest use of the model, setting regulatory standards, has a global use. Organizations like the European Union and the OECD, as well as many countries, apply SSD models to set (regulatory) standards. Those standards are then used prospectively, to evaluate whether the planned production, use or release of a (novel) chemical is sufficiently safe. If the predicted concentration exceeds the criterion, this is interpreted as a warning. Dependent on the regulatory context, the compound may be regulated, e.g., prohibited from use, or its use limited. The data used to build SSD models for deriving regulatory standards are often chronic test data, and no or low effect endpoints. The resulting standards have been evaluated in validation studies regarding the question of sufficient protection. Note that some jurisdictions have both protective standards as well as trigger values for remediation, based on SSD modelling.
2. The next use is in environmental quality assessment and management. In this case, the predicted or measured concentration of a chemical in an environmental compartment is often first compared to the reference level. This may already trigger management activities if the reference values have a regulatory status, such as a clean-up operation. The SSD may, however, be used to provide more detailed information on expected magnitude of impact, so that environmental management can prioritize most-affected sites for earlier remediation. The use of SSDs needs be tailored to the situation. That is, if the exposure concentrations form an array close to the reference value, the use of SSD_{NOEC}\s is a logical step, as this ranks the site pollution levels (via the PAFs) regarding the potentially affected fraction of species experiencing slight exceedances of the no effect level. If the study area contains highly polluted sites, that approach may show that all measured concentrations are in the upper tail of the SSD_{NOEC} sigmoid (horizontal part). In such cases, the SSD_{EC50} provides information on across-sites differences in expected impacts larger than the 50% effect level.
3. The third use is in Life Cycle Assessment of products. This use is comparative, so that consumers can select the most benign product, whilst producers can identify 'hot spots' of ecotoxicity in their production chains. A product often contains a suite of chemicals, so that the SSD model must be applied to all chemicals, by aggregating PAF-type outcomes over all chemicals. The model USEtox is the UN global consensus model for this application.

Today, these three forms of use of SSD models have an important role in the practice of environmental protection, assessment and management on the global scale, which relates to their intuitive meaning, their ease of use, and the availability of a vast number of ecotoxicity data in the global databases.

6.3.5. Question 1

What is the basic concept underlying SSD models?

6.3.5. Question 2

What is the main assumption underlying SSD models?

6.3.5. Question 3

What is meant by "the dual utility of an SSD model" in environmental protection, assessment and management?

6.3.5. Question 4

Given that the SSD model is a statistical description of ecotoxicological differences in sensitivity between species for a chemical, what is a critical step in the derivation and use of SSD model outputs?

6.3.5. Question 5

Does an SSD describe or explain differences in species sensitivity for a chemical?

6.3.6. Mixtures

under review

This page titled 2.3.6.3: Predictive risk assessment approaches and tools is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **6.3: Predictive risk assessment approaches and tools** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.6.4: Diagnostic risk assessment approaches and tools

6.4. Diagnostic risk assessment approaches and tools

Author: Michiel Kraak

Reviewers: Ad Ragas and Kees van Gestel

Learning objectives:

You should be able to

- define and distinguish hazard and risk
- distinguish predictive tools (toxicity tests) and diagnostic tools (bioassays)
- list bioassays and risk assessment tools at different levels of biological organization, ranging from laboratory to field approaches

Keywords: hazard assessment, risk assessment, prognosis, diagnosis, effect based monitoring, bioassays, effect directed analysis, mesocosm, biomonitoring, TRIAD approach, eco-epidemiology.

To determine whether organisms are at risk when exposed to certain concentrations of hazardous compounds in the field, the toxicity of environmental samples can be analysed. To this purpose, several approaches and techniques have been developed, known as *diagnostic* tools. The tools described in Sections 6.5.1-6.5.8 have in common that they make use of living organisms to assess environmental quality. This is generally achieved by performing bioassays in which the selected test species are exposed to (concentrates or dilutions of) environmental samples after which their performance (survival, growth, reproduction etc) is measured. The species selected as test organisms for bioassays are generally the same as the ones selected for toxicity tests (see section on Selection of ecotoxicity test organisms).

Each biological organization level has its own battery of test methods. At the lowest level of biological organization, a wide variety of *in vitro* bioassays is available (see section Effect based monitoring: *in vitro* bioassays). These comprise tests based on cell lines, but also bacteria and zebra fish embryos are employed. If the response of a bioassay to an environmental sample exceeds the predefined effect-based trigger value, the response is considered to be indicative of ecological risks. Yet, the compounds causing the observed toxicity are initially unknown. However, these can subsequently be elucidated with Effect Directed Analysis (see section Effect Directed Analysis). The sample causing the effect is subjected to fractionation and the fractions are tested again. This procedure is repeated until the sample is reduced to a few individual compounds, which can then be identified allowing to confirm their contribution to the observed toxic effects.

At higher levels of biological organization, a wide variety of *in vivo* tests and test organisms are available, including terrestrial and aquatic plants and animals (see section Effect based monitoring: *in vivo* bioassays). Yet, different test species tend to respond very differently to specific toxicants and specific field collected samples. Hence, the results of a single species bioassay may not reliably reflect the risk of exposure to a specific environmental sample. To avoid over- and underestimation of environmental risks, it is therefore advisable to employ a battery of *in vitro* and *in vivo* bioassays. In a case study on effect-based water quality assessment, we showed the great potential of this approach, resulting in the ranking of sites based on ecological risks rather than on the absence or presence of compounds (see section Effect based water quality assessment).

At the higher levels of biological organization, effect-based monitoring tools include bioassays performed in mesocosms (see section Community Ecotoxicology in practice) and in the field itself, the so called *in situ* bioassays (see section Biomonitoring: *in situ* bioassays and contaminant concentrations in organisms). Cosm studies represent a bridge between the laboratory and the natural world. The originality of mesocosms is based on the combination of ecological realism, the ability to manipulate different environmental parameters and still having the opportunity to replicate treatments.

In the field, the aim of biomonitoring is the *in situ* assessment of environmental quality on a regular basis in time, using living organisms (see section Biomonitoring: *in situ* bioassays and contaminant concentrations in organisms). Organisms are collected from reference sites and exposed in cages or artificial substrates at the study sites, after which they are recollected and either their condition is analysed (*in situ* bioassay) or the internal concentrations of specific target compounds are measured, or both (see section Biomonitoring: *in situ* bioassays and contaminant concentrations in organisms).

Finally, two approaches will be introduced that aid to bridge policy goals and ecosystem responses to perturbation: the TRIAD approach and eco-epidemiology. The TRIAD approach is a tool for site-specific ecological risk assessment, combining and integrating information on contaminant concentrations, bioassay results and ecological field inventories in a 'Weight of Evidence' approach (see section TRIAD approach). Eco-epidemiology is defined as the study of the distribution and causation of impacts of multiple stressor exposures in ecosystems, and the application of this study to reduce ecological impacts (see section Eco-epidemiology).

6.4. Question 1

Define and distinguish hazard and risk.

6.4. Question 2

Distinguish predictive tools (toxicity tests) and diagnostic tools (bioassays).

6.4. Question 3

What is the essence of diagnostic risk assessment?

6.4. Question 4

List and briefly describe bioassays at different levels of biological organisation.

6.4. Question 5

Name and briefly describe two approaches that aid to bridge policy goals and ecosystem responses to perturbation.

6.4.1. Effect-based monitoring: In vitro bioassays

Author: Timo Hamers

Reviewer: Beate Escher

Learning objectives:

You should be able to

- explain why effect-based monitoring is "more comprehensive" than chemical-analytical monitoring
- name several characteristics which make *in vitro* bioassays suitable for effect-based monitoring purposes
- give examples of most widely used bioassays
- describe the principles of a reporter gene assay, an enzyme induction assay, and an enzyme inhibition assay
- indicate how results from effect-based monitoring with *in vitro* bioassays can be interpreted in terms of environmental risk

Key words: effect-based monitoring; cell line; reporter gene assay; toxicity profile; trigger value

Effect-based monitoring

Diagnosis of the chemical status of the environment is traditionally performed by the analytical detection of a limited number of chemical compounds. Environmental quality is then assessed by making a compound-by-compound comparison between the measured concentration of an individual contaminant and its environmental quality standard (EQS). Such a compound-by-compound approach, however, cannot cover the full spectrum of contaminants given the unknown identity of the vast majority of compounds released into the environment. It also ignores the presence of unknown breakdown products formed during degradation processes and the presence of compounds with concentration levels below the analytical limit of detection. Furthermore, it overlooks combined effects of contaminants present in the complex environmental mixture.

To overcome these shortcomings, effect-based monitoring has been proposed as a comprehensive and cost-effective, complementary strategy to chemical analysis for the diagnosis of environmental chemical quality. In effect-based monitoring the toxic potency of the complex mixture is determined as a whole by testing environmental samples in bioassays. Bioassays are defined as "biological test systems that consist of whole organisms or parts of organisms (e.g., tissues, cells, proteins), which show a measurable and potentially biologically relevant response when exposed to natural or xenobiotic compounds, or complex mixtures present in environmental samples" (Hamers et al. 2010).

Bioassays making use of whole organisms are further referred to as *in vivo* bioassays (*in vivo* means "while living"). *In vivo* bioassays have relatively high ecological relevance as they provide information on survival, reproduction, growth, or behaviour of the species tested. *In vivo* bioassays will be addressed in a separate section.

In vitro bioassays

Bioassays making use of tissues, cells, proteins are called *in vitro* bioassays (*in vitro* means "in glass"), as - in the past -they were typically performed in test tubes or petri dishes made from glass. Nowadays, *in vitro* bioassays are more often performed in microtiter wells-plates containing multiple (6, 12, 24, 48, 96, 384, or 1536) test containers (called "wells") per plate. Most *in vitro* bioassays show a very mechanism-specific response, which is for instance indicative of the inhibition of a specific enzyme or the activation of a specific molecular receptor.

In addition to the mechanism-specific information about the complex mixture present in the environment, *in vitro* bioassays have several other advantages. Small test volumes, for instance, make the *in vitro* assays suitable to test small samples. If sampling volumes are not restricted, however, the small volume of the *in vitro* bioassays allow that pre-concentrated samples (i.e. extracts) can be tested. Moreover, *in vitro* bioassays have short test durations (usually incubation periods range from 15 minutes to 48 hours) and can be performed in relatively high-throughput, i.e. multiple samples can be tested per microtiter plate experiment. Microtiter plate experiments require an easy read-out (e.g. luminescence, fluorescence, optical density), which is typically a direct measure for the toxic potency to which the bioassay was exposed. Finally, using cells or proteins for toxicity testing raises less ethical objections than the use of intact organisms as is done in *in vivo* bioassays.

Cell-based *in vitro* bioassays can make use of different types of cells. Cells can be isolated from animal tissue and be grown in medium in cell culture flasks. If a flask grows full, cells can be diluted in fresh medium and be distributed over several new flasks (i.e. "passaging"). For cells freshly isolated from animal tissue (called primary cells), however, the number of passages is limited, due to the fact that the cells have a limited number of cell doublings. Thus, the use of primary cells in environmental monitoring is not preferred, as preparation of cell cultures is time-consuming and requires the use of animals. Moreover, the composition and activity of the cells may change from batch to batch. Instead, environmental monitoring often makes use of cell lines. A cell line is a cell culture derived from a single cell which has been immortalized, allowing the cell to divide infinitely. Immortalization of cells is obtained by selecting either a (mutated) cancer cell from a donor animal or human being, or by causing a mutation in an a healthy cell after isolation using chemicals or viruses. The advantage of a cell line is that all cells are genetically identical and can be used for an indefinite number of experiments. The drawback of cell lines is that the cells are cancer cells that do not behave like a healthy cell in an intact organism. For instance, cancer cells have lost their differentiated properties and have a short cell cycle due to increased proliferation (see section on *In vitro* toxicity testing).

Examples of

Reporter gene bioassays are a type of *in vitro* bioassays that are frequently used in effect-based monitoring. Such bioassays make use of genetically modified cell lines or bacteria that contain an incorporated gene construct encoding for an easily measurable protein (i.e. the reporter protein). This gene construct is developed in such a way that its expression is triggered by a specific interaction between the toxic compound and a cellular receptor. If the receptor is activated by the toxic compound, transcription and translation of the reporter protein takes place, which can be easily measured as a change in colour, fluorescence, or luminescence.

The most well-known reporter gene bioassays are steroid hormone-sensitive bioassays. These bioassays are based on the principle by which steroid hormones act, i.e. activation of a receptor protein followed by translocation of the hormone-receptor complex to the nucleus where it binds to a hormone-responsive element of the DNA, thereby initiating transcription and translation of steroid hormone-dependent genes. In case of a hormone-responsive reporter gene bioassay, the reporter gene construct is also under transcriptional control of a hormone-responsive element. Activation of the steroid hormone receptor by an endocrine disrupting compound thus leads to expression of the reporter protein, which can easily be measured. Estrogenic activity, for instance, is typically measured in cell lines in which a plasmid is stably transfected into the cellular genome that encodes for the reporter protein luciferase (Figure 1). Expression of this enzyme is under transcriptional control of an estrogen-responsive element (ERE). Upon exposure to an environmental sample, estrogenic compounds present in the sample may enter the cell and bind and activate the estrogen receptor (ER). The activated ER forms a dimer with another activated ER and is translocated to the nucleus where the dimer binds to the ERE, causing transcription and translation of the luciferase reporter gene. After 24 hours, the exposure is

terminated and the amount of luciferase enzyme can be easily quantified by lysis of the cells and adding the energy source ATP and the substrate luciferin. Luciferin is hydrolysed by luciferase, which is associated with the emission of light (i.e. the same reaction as occurs in fireflies or in glowing worms). The amount of light produced by the cells is quantified in a luminometer and is a direct measure for the estrogenic potency of the complex mixture to which the cells were exposed.

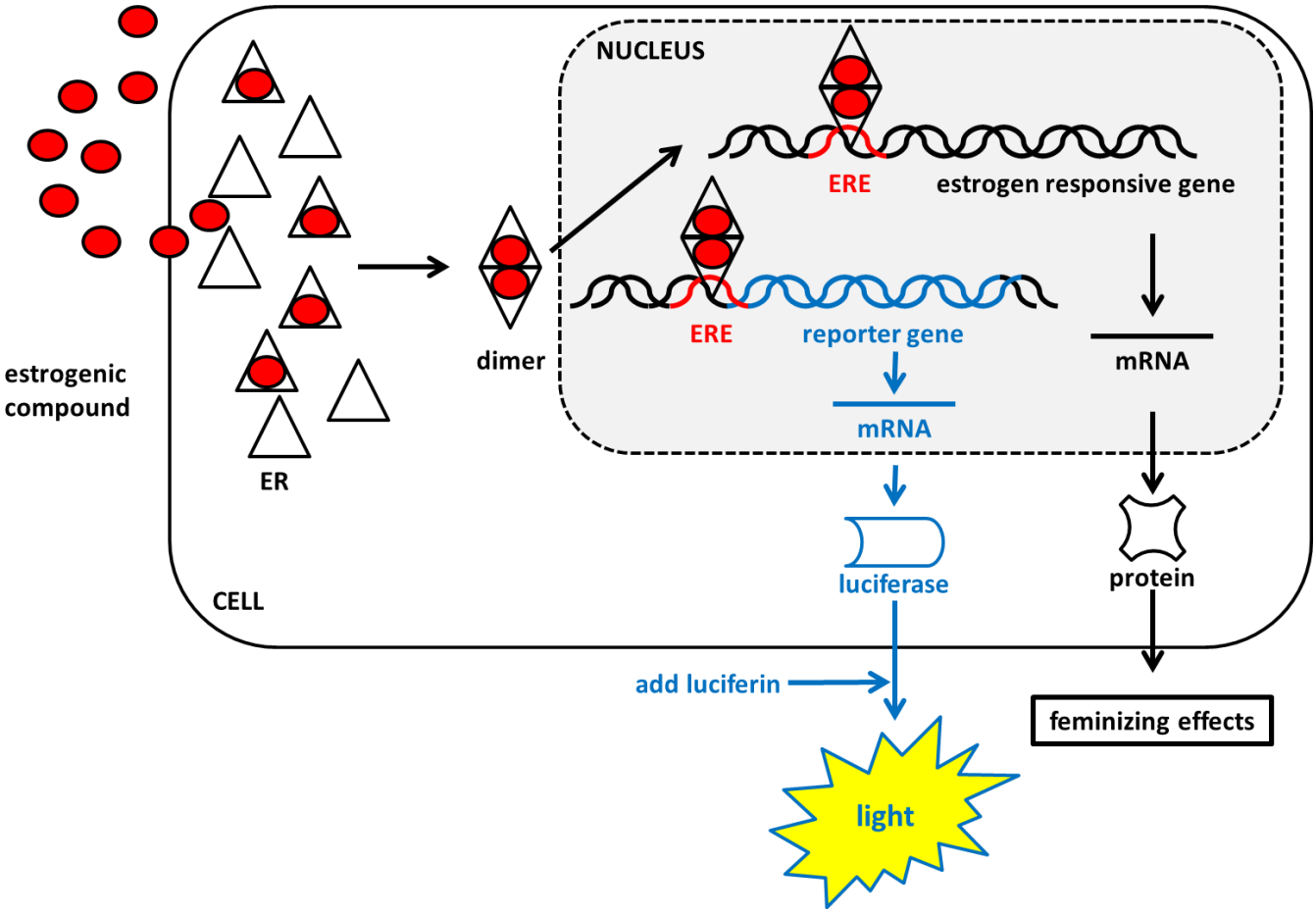


Figure 1: Principle of an estrogen responsive reporter gene assay: estrogenic compounds (red) enter the cell and activate the estrogen receptor (ER; triangle). Activated ERs form a dimer that is translocated to the nucleus where they bind to estrogen response elements (EREs). The regular subsequent pathway is indicated in black: estrogen responsive genes are transcribed into mRNA and translated into proteins that cause feminizing effects. The reporter gene pathway is indicated in blue: the reporter gene, which is also under transcriptional control of the ERE, is transcribed and translated into the reporter protein luciferase. Upon opening of the cell (lysis) and addition of the substrate luciferin and ATP as energy source, light is produced, which is a direct measure for the amount of luciferase produced, and thereby also for the estrogenic potency to which the cells were exposed.

Another classic bioassay for the detection of dioxin-like compounds is the ethoxyresorufin-o-deethylase (EROD) bioassay (Figure 2). The EROD bioassay is an enzyme induction bioassay that makes use of a hepatic cell line (i.e. derived from liver cells). Similar as described for the estrogenic compounds, dioxin-like compounds can enter these cells upon exposure to an environmental sample, and bind and activate a receptor protein, i.c. the arylhydrocarbon receptor (AhR) (see section on Receptor interactions). The activated AhR is subsequently translocated to the nucleus where it forms a dimer with another transcription factor (ARNT) that binds to the dioxin responsive element (DRE), causing transcription and translation of dioxin-responsive genes. One of these genes encodes for CYP1A1, a typical Phase I biotransformation enzyme. Upon lysis of the cells and addition of the substrate ethoxyresorufin, CYP1A1 is capable of hydrolysing this substrate into ethanol and resorufin, which is a fluorescent reaction product that can be measured easily. As such, the amount of fluorescence is a direct measure for the dioxin-like potency to which the cells were exposed.

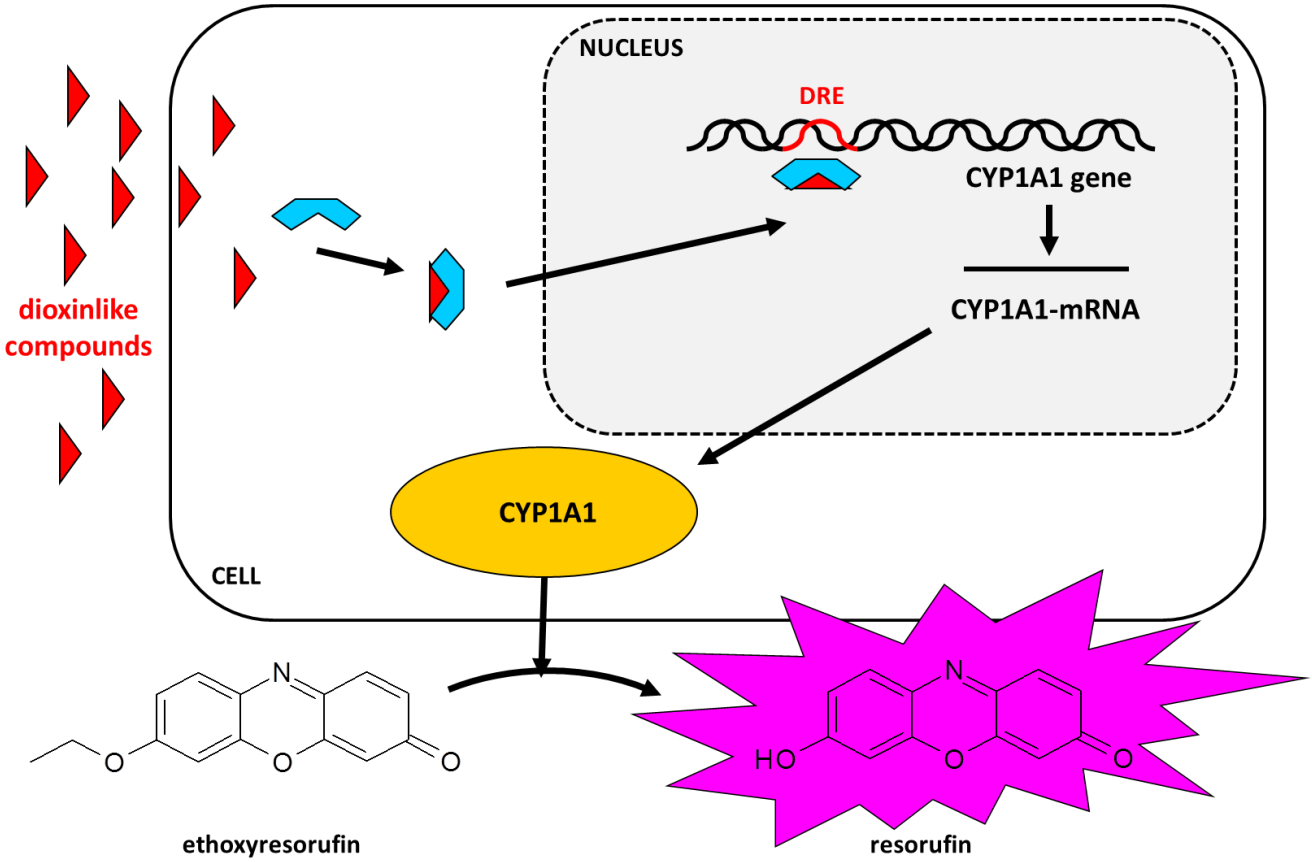


Figure 2: Simplified representation of the EROD (ethoxyresorufin-O-deethylase) assay. Dioxin-like compounds enter the hepatic cell and bind to the arylhydrocarbon receptor (AhR), which is translocated to the nucleus where it binds to dioxin-responsive elements (DREs) in the DNA. This causes transcription and translation of cytochrome P-4501A1 (CYP1A1). After 24h of incubation, the cells are lysed and the substrate ethoxyresorufin is added, which is oxidized by CYP1A1 into the fluorescent (pink) product resorufin

Another classic bioassay is the acetylcholinesterase (AChE) inhibition assay for the detection of organophosphate and carbamate insecticides (Figure 3). By making a covalent bond to the active site of the AChE enzyme, these compounds are capable of inhibiting the hydrolysis of the neurotransmitter acetylcholine (ACh) (see section on Protein inactivation). The *in vitro* AChE inhibition assay makes use of the principle that AChE can also hydrolyse an alternative substrate called acetylthiocholine (ATCh) into acetic acid and thiocholine (TCh). AChE inhibition leads to a decreased rate of TCh formation, which can be measured using an indicator, called Ellman's reagent. This indicator reacts with the thiol (-SH) group of TCh, resulting in a yellow breakdown product that can easily be measured photometrically. In the bioassay, purified AChE (commercially available for instance from electric eel) is incubated with an environmental sample in the presence of ATCh and Ellman's reagent. A decrease in the rate by which the yellow reaction product is formed is a direct measure for the inhibition of the AChE activity.

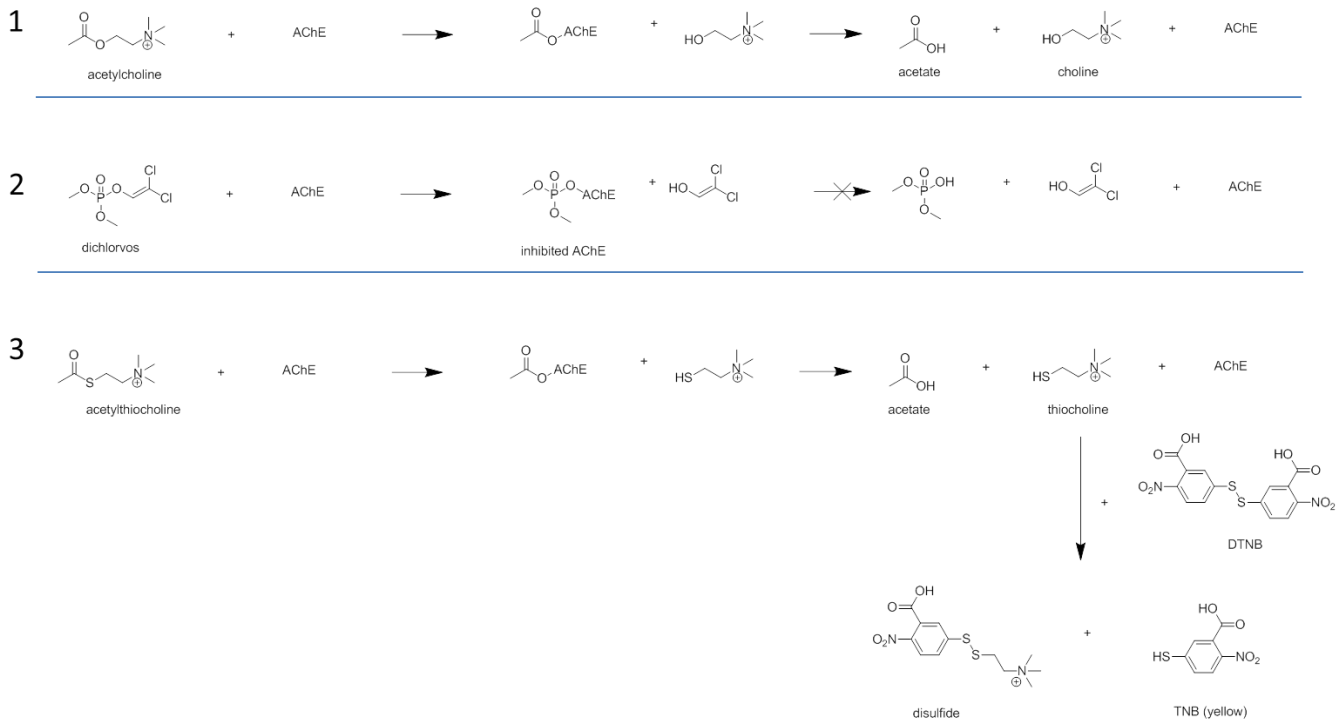


Figure 3: Principle of AChE inhibition: The normal hydrolysis of the neurotransmitter ACh by AChE is shown in the top row (1). The inhibition of AChE by the organophosphate insecticide dichlorvos is shown in the middle row (2). The phosphate ester-group does not release from the AChE active site, causing a decrease in AChE available for ACh hydrolysis. The principle of the AChE inhibition assay is shown in the bottom row (3). The remaining AChE activity is measured using an alternative substrate ATCh. The thiocholine product can be measured using the DTNB indicator (Ellman's reagent), which reacts with the thiol group, leading to a disulphide and a free TNB molecule. The yellow colour of the latter allows photometric quantification of the reaction.

Another bioassay that is used to detect mutagenic compounds in environmental samples is the Ames assay, which has been described in the section on Carcinogenicity and Genotoxicity.

Interpretation of the toxicity profile

In practice, multiple mechanism-specific *in vitro* bioassays are often combined into a test battery to cover the spectrum of toxicological endpoints in an (eco)system. As such, the battery can be considered as a safety net that signals the presence of toxic compounds at low concentrations. However, the question what combination of *in vitro* tests provides a sufficient level of coverage for the toxicological endpoints of concern still is an open one.

Still, testing an environmental sample in a battery of mechanism-specific *in vitro* bioassays yields a toxicity profile of the sample, indicating its toxic potency towards different endpoints. Two main strategies have been described to interpret *in vitro* toxicity profiles in terms of risk. In the "benchmark strategy", the toxicity profiles are compared to one or more reference profiles (Figure 4). A reference profile may be defined as the profile that is generally observed in environmental samples from locations with good chemical and/or ecological quality. The benchmark approach indicates to what extent the observed toxicity profile deviates from a toxicity profile corresponding to the desired environmental quality. It also indicates the endpoints that are most affected by the environmental sample.

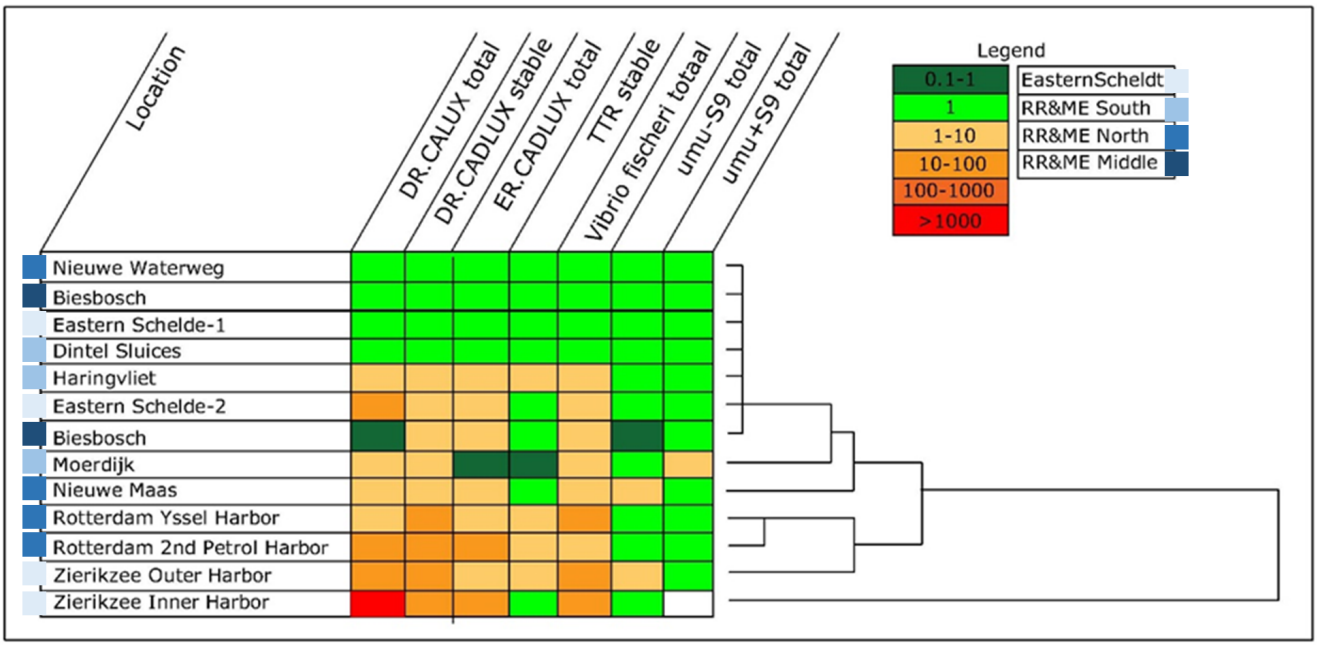


Figure. The tree-like structure (dendrogram) at the right indicates the relative distance between the different toxicity profiles. It clearly distinguishes between reference sites and clean sites on the one hand and harbour sites on the other hand. In between are samples from shipping lanes (Moerdijk and Nieuwe Maas). Zierikzee Inner Harbor is clearly a location with a deviating toxicity profile that is not similar to other harbour sites. Redrawn from Hamers et al. (2010) by Wilma IJzerman.

In the "trigger value strategy" the response of each individual bioassay is compared to a bioassay response level at which chemicals are not expected to cause adverse effects at higher levels of biological organization. This endpoint-specific "safe" bioassay response level is called an effect-based trigger (EBT) value. The method for deriving EBT values is still under development. It can be based on different criteria, such as laboratory toxicity data, field concentrations, or EU environmental quality standards (EQS) of individual compounds, which are translated into bioassay-specific effect-levels (see section on Effect-based water quality assessment).

In addition to the benchmark and trigger value approaches focusing on environmental risk assessment, effect-based monitoring with *in vitro* bioassays can also be used for effect-directed analysis (EDA). EDA focuses on samples that cause bioassay responses that cannot be explained by the chemicals that were analyzed in these samples. The goal of EDA is to detect and identify emerging contaminants that are responsible for the unexplained bioassay response and are not chemically analyzed because their presence or identity is unknown. In EDA, *in vitro* bioassay responses to fractionated samples are used to steer the chemical identification process of unknown compounds with toxic properties in the bioassays (see section on Effect-Directed Analysis).

Further reading:

Hamers, T., Leonards, P.E.G., Legler, J., Vethaak, A.D., Schipper, C.A. (2010). Toxicity profiling: an integrated effect-based tool for site-specific sediment quality assessment. Integrated Environmental Assessment and Management 6, 761-773

6.4.1. Question 1

Name advantages and disadvantages of effect-based and chemical-based monitoring strategies?

6.4.1. Question 2

Name at least three characteristics that make *in vitro* bioassays suitable for effect-based monitoring.

6.4.1. Question 3

Can the principle of the EROD assay also be used to develop a reporter gene assay? Explain your answer.

6.4.1. Question 4

In the benchmark approach (see text), toxicity profiles from sampling locations are compared to a reference profile. Should the reference profile always correspond to a clean situation? Motivate your answer.

6.4.2. Effect Directed Analysis

Author: Marja Lamoree

Reviewers: Timo Hamers, Jana Weiss

Learning goals:

You should be able to

- explain the complementary nature of the analytical/chemical and biological/toxicological techniques used in Effect-Directed Analysis
- explain the purpose of Effect-Directed Analysis
- describe the steps in the Effect-Directed Analysis process
- describe when the application of Effect-Directed Analysis is most useful

Keywords: extraction, bioassay testing, fractionation, identification, confirmation

In general, the quality of the environment may be monitored by two complementary approaches: i) quantitative chemical analysis of selected (priority) pollutants and ii) effect-based monitoring using in vitro/vivo bioassays. Compared to the more classical chemical analytical approach that has been used for decades, effect-based monitoring is currently applied in an explorative manner and has not yet matured into a routinely implemented monitoring tool that is anchored in legislation. However, in an international framework, developments to formalize the role of effect-based monitoring and to standardize the use of bioassay testing for environmental quality assessment are underway.

A weakness of the chemical approach is that because of the preselection of target compounds for quantitative analysis other compounds that are of relevance for the environmental quality may be missed. In comparison, inclusiveness is one of the advantages of effect-based monitoring: all compounds - and not only a few pre-defined ones - having a specific effect will contribute to the total, measured biological activity (see Section In vitro bioassays). In turn, the effect-based approach strongly benefits from chemical analytical support to pinpoint which compounds are responsible for the observed activity and to be able to take measures for environmental protection, e.g. the reduction of the emission or discharge of a specific toxic compound into the environment.

In Effect-Directed Analysis (EDA), the strengths of analytical chemical techniques and effect-based testing are combined with the aim to identify novel compounds that show activity in a biological analysis and that would have gone unnoticed using the chemical and the effect-based approach separately. A schematic representation of EDA is shown in Figure 1 and the various steps are described below in more detail. There is no limitation regarding the sample matrix: EDA has been applied to e.g. water, soil/sediment and biota samples. It is used for in-depth investigations at locations that are suspected to be contaminated but where the compounds responsible for the observed adverse effects are not known. In addition to environmental quality assessment, EDA is applied in the fields of food security analysis and drug discovery. In Table 1 examples of EDA studies are given.

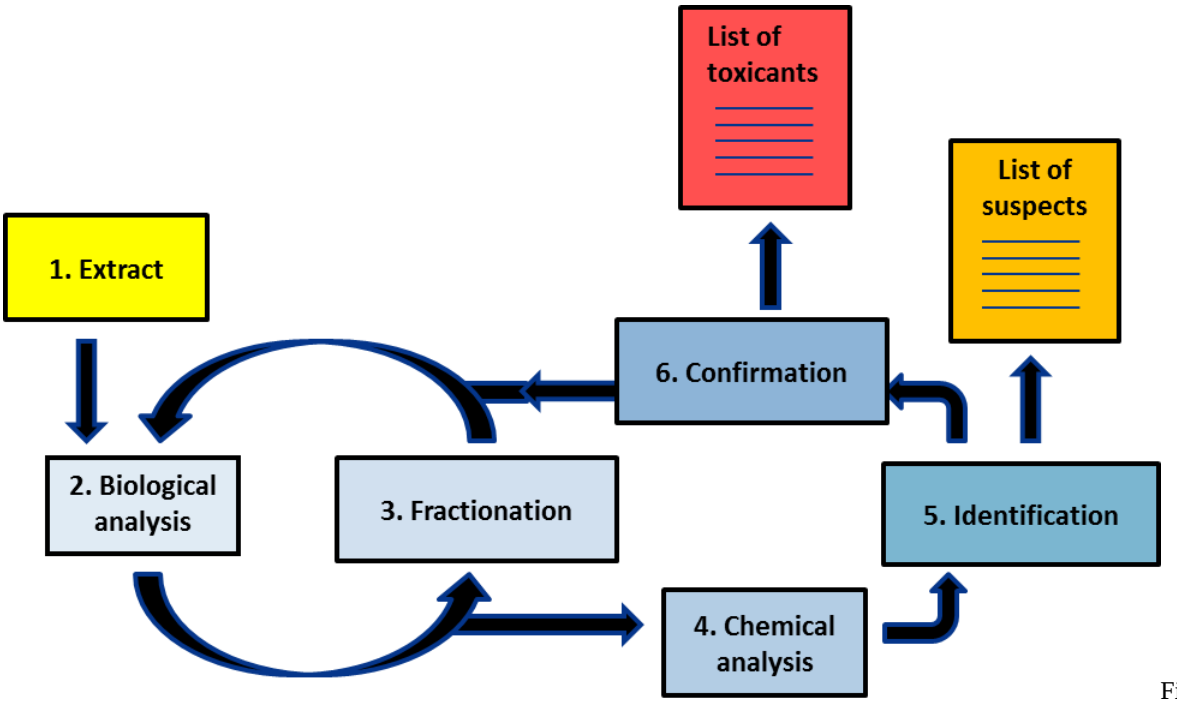


Figure 1. Schematic representation of Effect-Directed Analysis (EDA).

1. Extract

The first step is the preparation of an extract of the sample. For soil/sediment samples, a sieving step prior to the actual extraction may be necessary in order to remove large particles and obtain a sample that is well-defined in terms of particle size (e.g. <200 µm). Examples of biota samples are whole organism homogenates or parts of the organism, such as blood and liver. For the extraction of the samples, analytical techniques such as liquid/liquid or solid phase extraction are applied to concentrate the compounds of interest and to remove matrix constituents that may interfere with the later steps of the EDA.

2. Biological analysis

The choice of endpoint to include in an EDA study is very important, as it dictates the nature of the toxicity of the compounds that may be identified (see Section on Toxicodynamics and Molecular Interaction). For application in EDA, typically *in vitro* bioassays that are carried out in multiwell (≥ 96 well) plates can be used, because of their low cost, high throughput and ease of use (see Section on In vitro bioassays), although sometimes *in vivo* assays (see Section on In vivo bioassays) are applied too.

Table 1. Examples of EDA studies, including endpoint, type of bioassay, sample matrix and compounds identified.

Endpoint	Type of bioassay	Sample matrix	Type of compounds identified
	In vitro		
Estrogenicity	Cell based reporter gene	Sediment	Endogenic hormones
Anti-androgenicity	Cell based reporter gene	Sediment	Plasticizers, organophosphorus flame retardants, synthetic fragrances
	idem	Water	Pharmaceuticals, pesticides, plasticizers, flame retardants, UV filters
Mutagenicity	Bacterial luminescence reporter strain	Water	Benzotriazoles
Thyroid hormone disruption	Radioligand binding	Polar bear plasma	Metabolites of PCBs, nonylphenols
	In vivo		
Photosystem II toxicity	Pulse Amplitude Modulation fluorometry	Water	Pesticides
Endocrine disruption	Snail reproduction	Sediment	Phthalates, synthetic fragrances, alkylphenols

3. Fractionation

Fractionation of the extract is achieved by the application of chromatography, resulting in the separation of the - in most cases - multitude of different compounds that are present in an extract of an environmental sample. Chromatographic separation is obtained after the migration of compounds through a sorbent bed. In most cases, the separation principle is based on the distribution of compounds between the liquid mobile phase and the solid stationary phase (liquid chromatography, or LC), but a chromatographic separation using the partitioning between the gas phase and a sorbent bed (gas chromatography, or GC) is also possible. At the end of the separation column, at specified time intervals fractions can be collected that are simpler in composition in comparison to the original extract: a reduction in the number of compounds per fraction is obtained. The collected fractions are tested in the bioassay and the responsive fractions are selected for further chemical analysis and identification (step 4). The time intervals for fraction collection vary between a few minutes in older applications and a few seconds in new applications of EDA, which enables fractionation directly into multiwell plates for high throughput bioassay testing. In cases where fractions are collected during time intervals in the order of minutes, the fractions are still so complex that a second round of fractionation to obtain fractions of reduced complexity is often necessary for the identification of compounds that are responsible for the observed effect (see Figure 2).

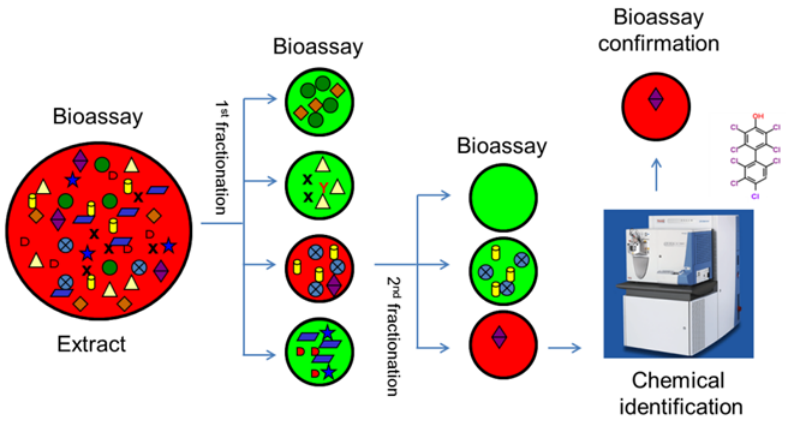


Figure 2. Schematic representation of extract fractionation and selection of fractions for further testing, identification and confirmation.

4. Chemical Analysis

Chemical analysis for the identification of the compounds that cause the effect in the bioassay is usually done by LC coupled to mass spectrometric (MS) detection. To obtain high mass accuracy that facilitates compound identification, high resolution mass spectrometry (HR-MS) is generally applied. Fractions obtained after one or two fractionation steps are injected into the LC-MS system. In studies where fractionation into multiwell plates is used (and thus small fractions in the order of microliters are collected), only one round of fractionation is applied. In these cases, identification and fraction collection can be done in parallel, using a splitter after the chromatographic column that directs part of the eluent from the column to the well plate and the other part to the MS (see Figure 3). This is called high throughput EDA (HT-EDA).

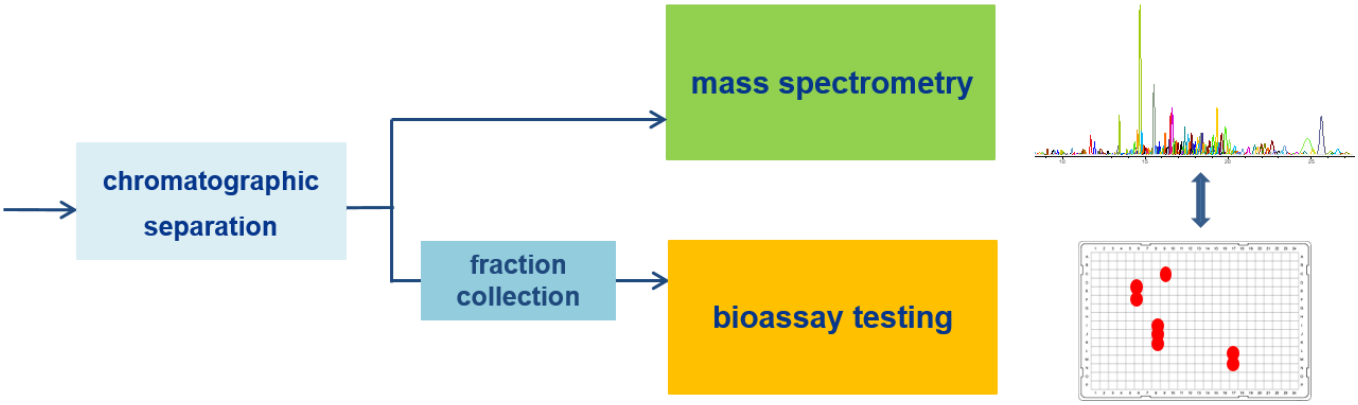


Figure 3. Schematic representation of high-throughput Effect-Directed Analysis (HT-EDA).

5. Identification

The use of HR-MS is necessary to obtain mass information to establish the molecular weight with high accuracy (e.g. 119.12423 Dalton) to derive the molecular formula (e.g. $C_6H_5N_3$) of the compound. Optimally, HR-MS instrumentation is equipped with an MS-MS mode, in which compound fragmentation is induced by collisions with other molecules, resulting in fragments that are specific for the original compound. Fragmentation spectra obtained using the MS-MS mode of HR-MS instruments help to elucidate the structure of the compounds eluting from the column, see for an example Figure 4.

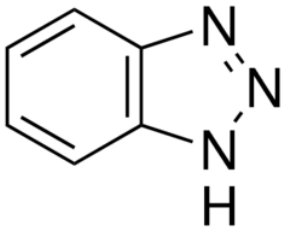


Figure 4. Example of a chemical structure corresponding to an accurate mass of 119.12423 Dalton and the corresponding molecular formula $C_6H_5N_3$: 1,2,3-benzotriazole.

Other information such as log K_{ow} may be calculated using dedicated software packages that use elemental composition and structure as input. To aid the identification process, compound and mass spectral libraries are used as well as the more novel databases containing toxicity information (e.g. PubChem Bioassay, Toxcast). Mass spectrometry instrumentation vendor software, public/web-based databases and databases compiled in-house enable suspect screening to identify compounds that are known, e.g. because they are applied in consumer products or construction materials. When MS signals cannot be attributed to known compounds or their metabolites/transformation products, the identification approach is called non-target screening, where additional identification techniques such as Nuclear Magnetic Resonance (NMR) may aid the identification. The identification process is complicated and often time consuming, and results in a suspect list that needs to be evaluated for further confirmation of the identification.

6. Confirmation

For an unequivocal confirmation of the identity of a tentatively identified compound, it is necessary to obtain a standard of the compound to investigate whether its analytical chemical behaviour corresponds to that of the tentatively identified compound in the environmental sample. In addition, the biological activity of the standard should be measured and compared with the earlier obtained data. In case both the chemical analysis and bioassay testing results support the identification, confirmation of compound identity is achieved.

In principle, the confirmation step of an EDA study is very straightforward, but in current practice the standards are mostly not commercially available. Dedicated synthesis is time consuming and costly, therefore the confirmation step often is a bottleneck in EDA studies.

The application of EDA is suitable for samples collected at specific locations where comprehensive chemical analysis of priority pollutants and other chemicals of relevance has been conducted already, and where ecological quality assessment has revealed that the local conditions are compromised (see other Sections on Diagnostic risk assessment approaches and tools). Especially those samples that show a significant difference between the observed (in vitro) bioassay response and the activity that may be calculated according to the concept of Concentration Addition (see Section on Mixture Toxicity) by using the relative potencies and the concentrations of compounds active in that bioassay need a further in-depth investigation. EDA can be implemented at these 'hotspots' of environmental contamination to unravel the identity of compounds that have an effect, but that were not included in the chemical monitoring of the environmental quality. Knowledge with regard to the main drivers of toxicity at a specific location supports accurate decision making that is necessary for environmental quality protection.

6.4.2. Question 1

Draw a scheme of EDA and name the different steps.

6.4.2. Question 2

What is the aim of the fractionation of an extract?

6.4.2. Question 3

Describe the confirmation step of EDA.

6.4.2. Question 4

Give an example of an EDA study with regard to endpoint, bioassay, matrix and type of compounds identified.

6.4.2. Question 5

Explain why quantitative chemical analysis of known pollutants of e.g. a water sample is complementary to effect-based testing of that sample using e.g. an in vitro bioassay?

6.4.3. Effect-based monitoring: In vivo bioassays

Effect based monitoring: *in vivo* bioassays

Authors: Michiel Kraak, Carlos Barata

Reviewers: Kees van Gestel, Jörg Römbke

Learning objectives:

You should be able to:

- define *in vivo* bioassays and to explain how *in vivo* bioassays are performed.
- give examples of the most commonly used *in vivo* bioassays per environmental compartment.
- motivate the necessity to incorporate several *in vivo* bioassays into a bioassay battery.

Key words: risk assessment, diagnosis, effect based monitoring, *in vivo* bioassays, environmental compartment, bioassay battery

Introduction

To determine whether organisms are at risk when exposed to hazardous compounds present at contaminated field sites, the toxicity of environmental samples can be analysed. To this purpose, several diagnostic tools have been developed, including a wide variety of in *vitro*, *in vivo* and *in situ* bioassays (see sections on *In vitro* bioassays and on In situ bioassays). *In vivo* bioassays make use of whole organisms (*in vivo* means "while living"). The species selected as test organisms for *in vivo* bioassays are generally the same as the ones selected for single species toxicity tests (see sections 4.3.4, 4.3.5, 4.3.6 and 4.3.7 on the Selection of ecotoxicity test organisms). Likewise, also the endpoints measured in *in vivo* bioassays are the same as those in single species ecotoxicity tests (see section on Endpoints). *In vivo* bioassays therefore have a relatively high ecological relevance, as they provide information on the survival, reproduction, growth, or behaviour of the species tested. A major difference between toxicity tests and bioassays is the selection of the controls. In laboratory toxicity experiments the controls consist of non-spiked 'clean' test medium (see section on Concentration response relationships). In bioassays the choice of the controls is more complicated though. Non-treated test medium may be incorporated as a control in bioassays to check for the health and quality of the test organisms. But control media, like standard test water or artificial soil and sediment may differ in numerous aspects from natural environmental samples. Therefore, the control should preferably be a test medium that has exactly the same physicochemical properties as the contaminated sample, except for the chemical pollutants being present. This ideal situation, however, hardly ever exists. Hence, it is recommended to also incorporate environmental samples from less or non-contaminated reference sites into the bioassay and to compare the response of the organism to samples from contaminated sites with those from reference sites. Alternatively, controls can be selected as the least contaminated environmental samples from a gradient of pollution or as the dilution required to obtain no effect. As dilution medium artificial control medium can be used or medium from a reference site.

The most commonly used

For the soil compartment, the earthworms *Eisenia fetida*, *E. andrei* and *Lumbricus rubellus*, the enchytraeid *Enchytraeus crypticus* and the collembolan *Folsomia candida* are most frequently selected as *in vivo* bioassay test organisms. An example of employing earthworms to assess the ecotoxicological effects of Pb contaminated soils is given in Figure 1. The figure shows the total Pb concentrations in different field soils taken from a soccer field (S), a bullet plot (B), grassland (G1, G3) and a forest (F1-F3) site near a shooting range. The pH of the grassland soils was near neutral (pH_{CaCl2} = 6.5-6.8), but the pH was rather low (3.2-3.7) for all other field sites. Earthworms exposed to these soils showed a significantly reduced reproductive output (Figure 1) at the most contaminated sites. At the less contaminated sites, earthworm responses were also affected by the difference in soil pH, leading to low juvenile numbers in the acid soil F0 but high numbers in the near neutral reference R3 and the field soil G3. In fact, earthworm reproduction was highest in the latter soil, even though it did contain an elevated concentration of 355 ± 54 mg Pb/kg dry soil. In soil G1, which contained almost twice as much Pb (656 ± 60 mg Pb/kg dry soil), reproduction was much lower and also reduced compared to the control, suggesting the presence of additional, unknown stressor (Luo et al., 2014).

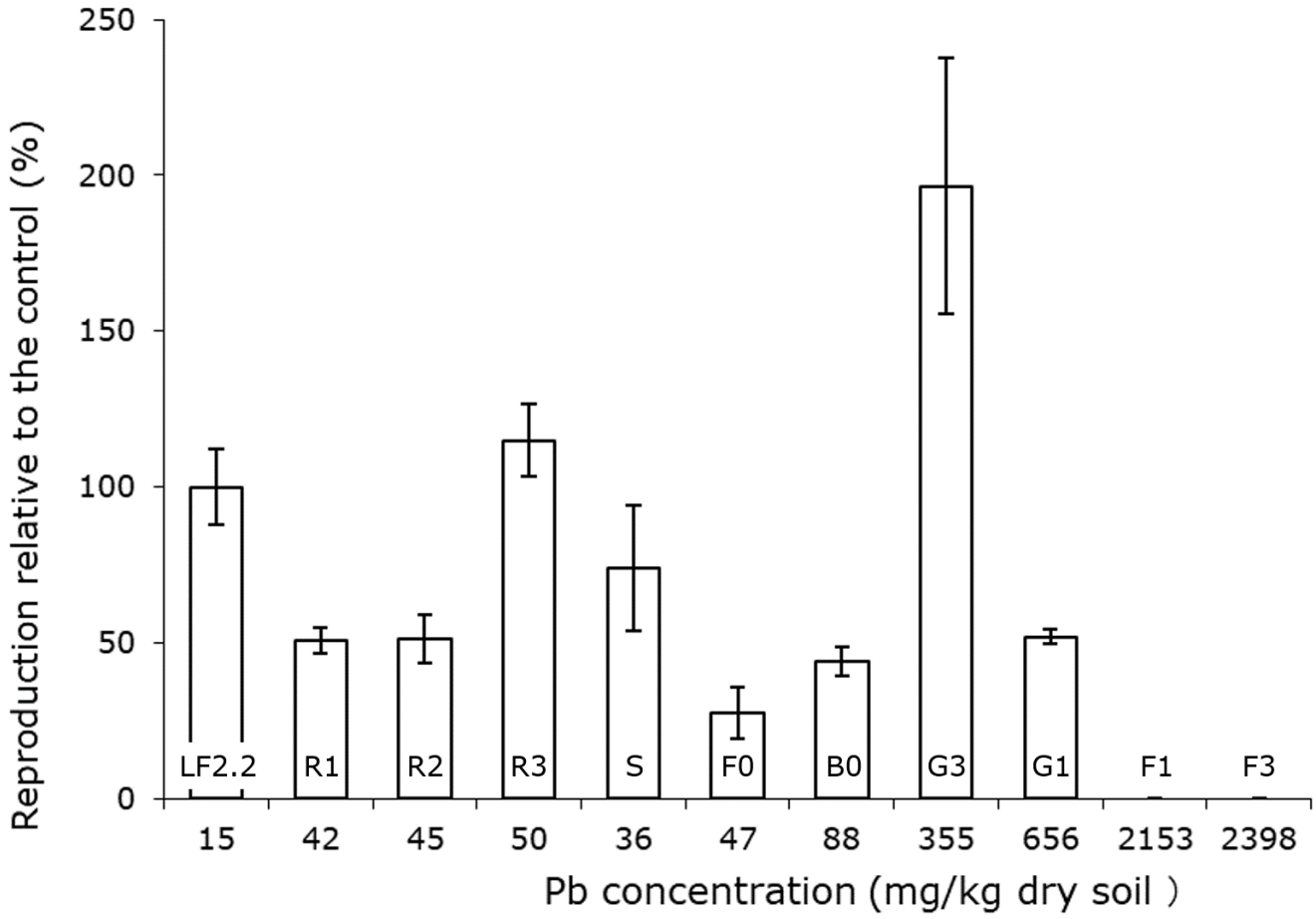


Figure 1. Reproduction of the earthworm Eisenia andrei after 4 weeks of exposure to control soils (LF2.2, R1, R2, R3) and field soils (S, B0, G1, G3, F0, F1, F3) from a Pb pollution gradient near a shooting range. Shown are the mean relative numbers of juveniles ± SD (n=4-5), compared to the control Lufa 2.2 (LF2.2) soil, as a function of average total Pb concentrations in the soils. Data from Luo et al. (2014).

For water, predominantly daphnids are employed, mainly *Daphnia magna*, but sometimes also other daphnid species or other aquatic invertebrates are selected. Also bioassays with several primary producers are available. An example of exposing daphnids (*Chydorus sphaericus*) to water samples is shown in Figure 2. The bars show the toxicity of the water samples and the diamonds the concentrations of cholinesterase inhibitors, as a proxy for the presence of insecticides. The toxicity of the water samples was higher when also the concentrations of insecticides were higher. Hence, in this case, the observed toxicity is well explained by the measured compounds. Yet, it has to be realized that this is an exception rather than a rule, since mostly a large portion of the toxic effects observed in surface waters cannot be attributed to compounds measured by water authorities and moreover, interactions are also not covered by such analytical data (see section on Effect based water quality assessment).

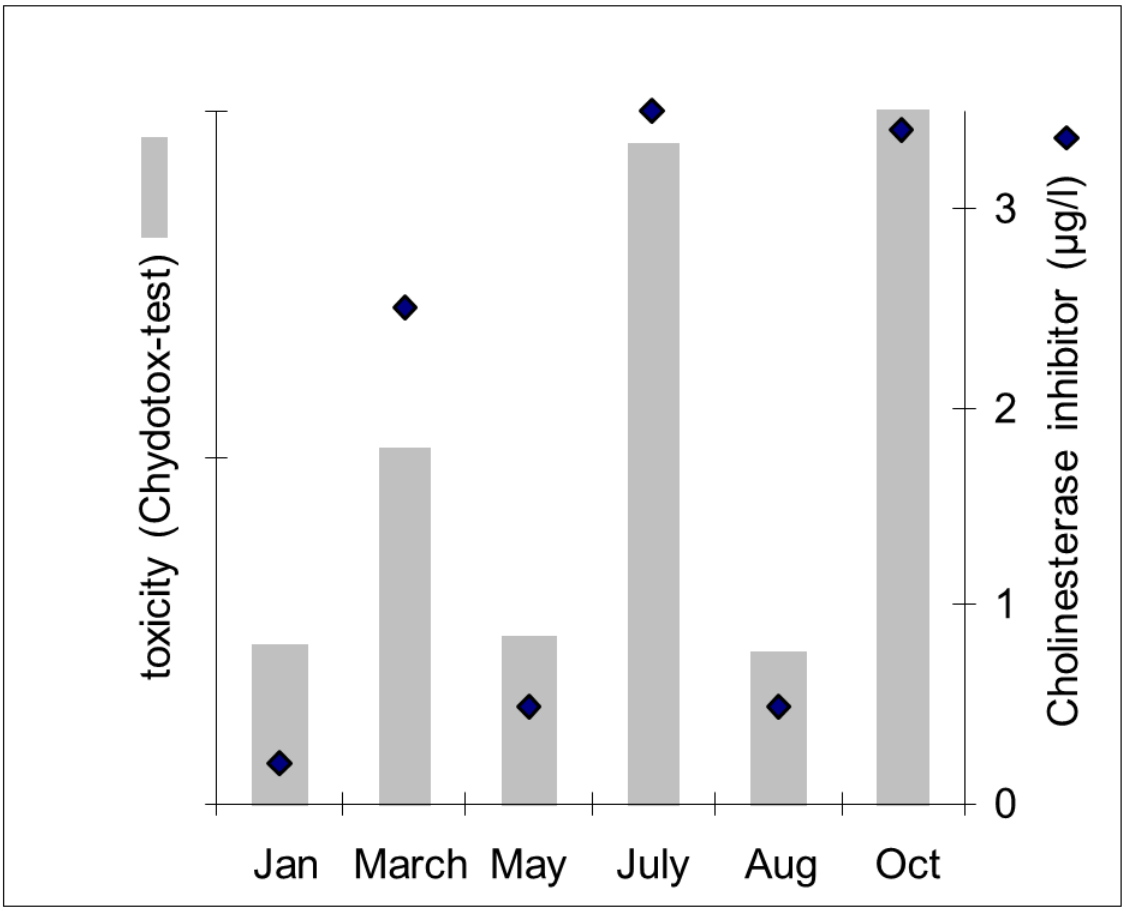


Figure 2.
Toxicity of water samples to daphnids (Chydorus sphaericus)(bars) and the concentrations of cholinesterase inhibitors, as a proxy for the presence of insecticides (diamonds). Data from Pieters et al. (2008).

For sediments, oligochaetes and chironomids are selected as test organisms, but sometimes also rooting macrophytes and benthic diatoms. An example of exposing chironomids (*Chironomus riparius*) to contaminated sediments is shown in Figure 3. Whole sediment bioassays with chironomids allow the assessment of sensitive species-specific sublethal endpoints (see section on Chronic toxicity), in this case emergence. Figure 3 shows that more chironomids emerged on the reference sediment than on the contaminated sediment and that the chironomids on the reference sediment also emerged faster than on the contaminated sediment.

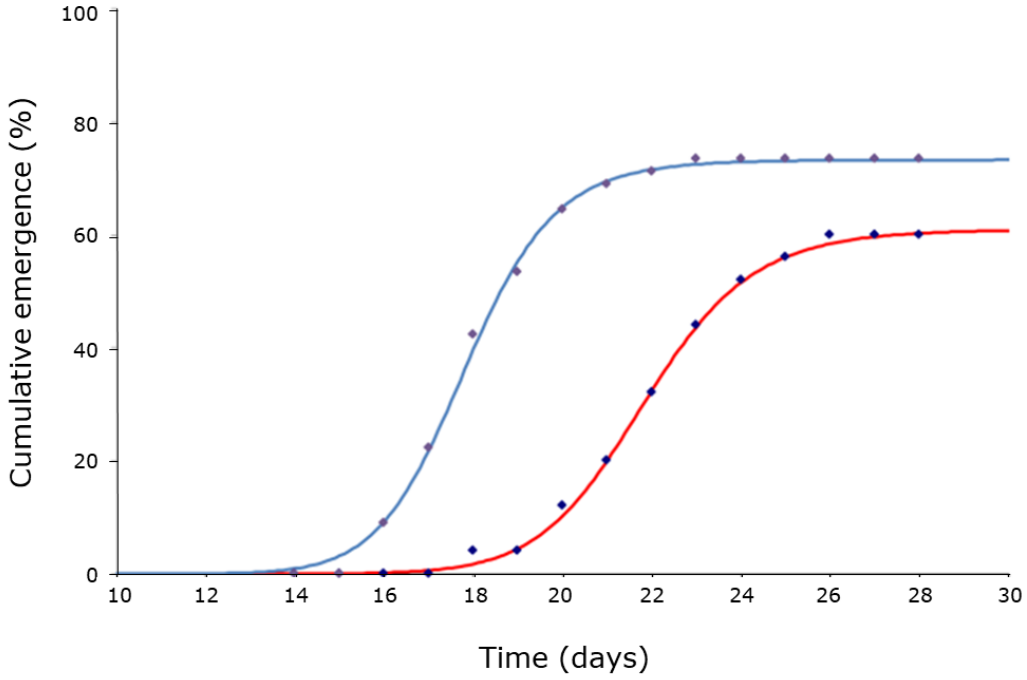


Figure 3. Emergence of chironomids (Chironomus riparius) on a reference (blue line) and a contaminated sediment (red line). Data from Nienke Wieringa.

For sediment, also benthic diatoms are selected as *in vivo* bioassay test organisms. Figure 4 shows the growth of the benthic diatom *Nitzschia perminuta* after 4 days of exposure to 160 sediment samples. The dotted line represents control growth. The growth of the diatoms ranged from higher than the controls to no growth at all, raising the question which deviation from the control should be considered a significant adverse effect.

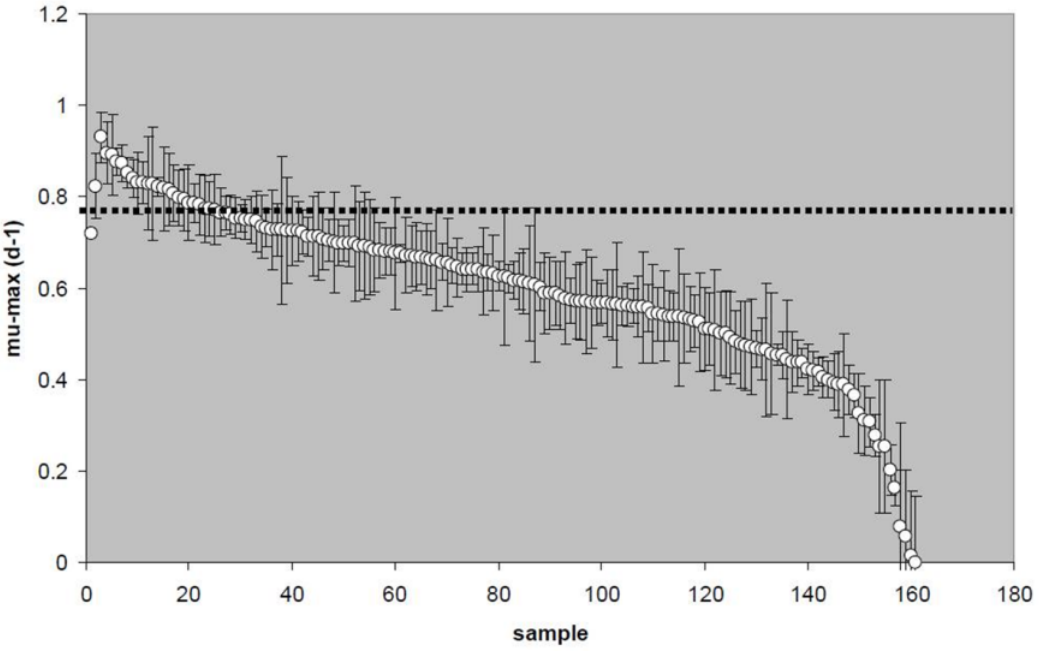


Figure 4. Growth of the benthic diatom Nitzschia perminuta after 4 days of exposure to 160 sediment samples. The dotted line represents control growth. Data from Harm van der Geest.

In vivo bioassay batteries

Environmental quality assessments are often performed with a single test species, like the four examples given above. Yet, toxicity is species and compound specific and this may therefore result in large margins of uncertainty in the environmental quality assessments, consequently leading to over- or underestimation of environmental risks. Obvious examples include the presence of herbicides that only would induce responses in bioassays with primary producers and the other way around, the presence of insecticides that induces strong effects on insects and to a lesser extent on other animals, but that would be completely overlooked in bioassays with primary producers. To reduce these uncertainties and to increase ecological relevance it is therefore advised to incorporate more test species belonging to different taxa in a bioassay battery (see section on Effect based water quality assessment).

References

Luo, W., Verweij, R.A., Van Gestel, C.A.M. (2014). Determining the bioavailability and toxicity of lead to earthworms in shooting range soils using a combination of physicochemical and biological assays. Environmental Pollution 185, 1-9.

Pieters, B.J., Bosman-Meijerman, D., Steenbergen, E., Van den Brandhof, E.-J., Van Beelen, P., Van der Grinten, E., Verweij, W., Kraak, M.H.S. (2008). Ecological quality assessment of Dutch surface waters using a new bioassay with the cladoceran *Chydorus sphaericus*. Proceedings Netherlands Entomological Society Meetings 19, 157-164.

6.4.3. Question 1

Define *in vivo* bioassays and explain how *in vivo* bioassays are performed.

6.4.3. Question 2

Give examples of the most commonly used *in vivo* bioassays per environmental compartment.

6.4.3. Question 3

Motivate the necessity to incorporate several *in vivo* bioassays into a bioassay battery.

6.4.4. Effect Based water quality assessment

Effect-based water quality assessment

Authors: Milo de Baat, Michiel Kraak

Reviewers: Ad Ragas, Ron van der Oost, Beate Escher

Learning objectives:

You should be able to

- list the advantages and drawbacks of an effect-based monitoring approach in comparison to a compound-based approach for water quality assessment.
- motivate the necessity of employing a bioassay battery in effect-based monitoring approaches.
- explain the expression of bioassay responses in terms of toxic/bioanalytical equivalents of reference compounds.
- translate the outcome of a bioassay battery into a ranking of contaminated sites based on ecotoxicological risk.

Keywords: Effect-based monitoring, water quality assessment, bioassay battery, effect-based trigger values, ecotoxicological risk assessment

Introduction

Traditional chemical water quality assessment is based on the analysis of a list of a varying, but limited number of priority substances. Nowadays, the use of many of these compounds is restricted or banned, and concentrations of priority substances in surface waters are therefore decreasing. At the same time, industries have switched to a plethora of alternative compounds, which may enter the aquatic environment, seriously impacting water quality. Hence, priority substances lists are outdated, as the selected compounds are frequently absent, while many compounds with higher relevance are not listed as priority substances. Consequently, a large portion of toxic effects observed in surface waters cannot be attributed to compounds measured by water authorities, and toxic risks to freshwater ecosystems are thus caused by mixtures of a myriad of (un)known, unregulated compounds. Understanding of these risks requires a paradigm shift towards new monitoring methods that do not depend on chemical analysis of

priority substances solely, but consider the biological effects of the entire micropollutant mixture first. Therefore, there is a need for effect-based monitoring strategies that employ bioassays to identify environmental risk. Responses in bioassays are caused by all bioavailable (un)known compounds and their metabolites, whether or not they are listed as priority substances.

Table 1. Example of the bioassay battery employed by the SIMONI approach of Van der Oost et al. (2017) that can be applied to assess surface water toxicity. Effect-based trigger values (EBT) were previously defined by Escher et al. (2018) (PAH, anti-AR and ER CALUX) and Van der Oost et al. (2017).

Bioassay battery

The regular application of effect-based monitoring largely relies on the ease of use, endpoint specificity, costs and size of the used bioassays, as well as on the ability to interpret the measured responses. To ensure sensitivity to a wide range of potential stressors, while still providing specific endpoint sensitivity, a successful bioassay battery like the example given in Table 1 can include *in situ* whole organism assays (see section on Biomonitoring and in situ bioassays), and should include laboratory-based whole-organism *in vivo* (see section on In vivo bioassays) and mechanism-specific *in vitro* assays (see section on In vitro bioassays). Adverse effects in the whole-organism bioassays point to general toxic pressure and represent a high ecological relevance. *In vitro* or small-scale *in vivo* assays with specific drivers of adverse effects allow for focused identification and subsequent confirmation of (groups of) toxic compounds with specific modes of action. Bioassay selection can also be based on the Adverse Outcome Pathways (AOP) (see section on Adverse Outcome Pathways) concept that describes relationships between molecular initiating events and adverse outcomes. Combining different types of bioassays ranging from whole organism tests to *in vitro* assays targeting specific modes of action can thus greatly aid in narrowing down the number of candidate compound(s) that cause environmental risks. For example, if bioanalytical responses at a higher organisational level are observed (the orange and black pathways in Figure 1), responses in specific molecular pathways (blue, green, grey and red in Figure 1) can help to identify certain (groups of) compounds responsible for the observed effects.

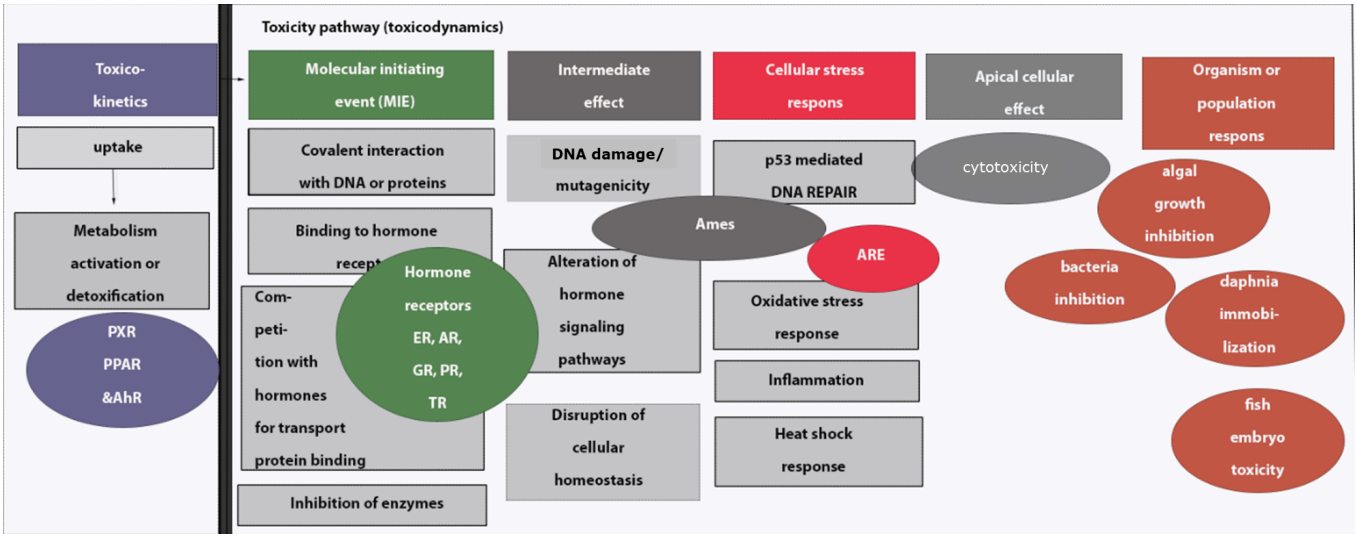


Figure 1. From toxicokinetics via molecular responses to population responses. Redrawn from Escher et al. (2018) by Wilma IJzerman.

Toxic and bioanalytical equivalent concentrations

The severity of the adverse effect of an environmental sample in a bioassay is expressed as toxic equivalent (TEQ) concentrations for toxicity in *in vivo* assays or as bioanalytical equivalent (BEQ) concentrations for responses in *in vitro* bioassays. The toxic equivalent concentrations and bioanalytical equivalent concentrations represent the joint toxic potency of all unknown chemicals present in the sample that have the same mode of action (see section on Toxicodynamics and molecular interactions) as the reference compound and act concentration-additively (see section on Mixture toxicity). The toxic equivalent concentrations and bioanalytical equivalent concentrations are expressed as the concentration of a reference compound that causes an effect equal to the entire mixture of compounds present in an environmental sample. Figure 2 depicts a typical dose-response curve for a molecular *in vitro* assay that is indicative of the presence of compounds with a specific mode of action targeted by this *in vitro*

	Bioassay	Endpoint	Reference compound	EBT	Unit
in situ	Daphnia in situ	Mortality	n/a	20	% mortality
in vivo	Daphniatox	Mortality	n/a	0.05	TU
	Algatox	Algal growth inhibition	n/a	0.05	TU
	Microtox	Luminescence inhibition	n/a	0.05	TU
in vitro CALUX	cytotox	Cytotoxicity	n/a	0.05	TU
	DR	Dioxin (-like) activity	2,3,7,8-TCDD	50	pg TEQ/L
	PAH	PAH activity	benzo(a)pyrene	6.21	ng BapEQ/L
	PPARγ	Lipid metabolism inhibition	rosiglitazone	10	ng RosEQ/L
	Nrf2	Oxidative stress	curcumin	10	µg CurEQ/L
	PXR	Toxic compound metabolism	nicardipine	3	µg NicEQ/L
	p53 -S9	Genotoxicity	n/a	0.005	TU
	p53 +S9	Genotoxicity (after metabolism)	n/a	0.005	TU
	ER	Estrogenic activity	17ß-estradiol	0.1	ng EEQ/L
	anti-AR	Antiandrogenic activity	flutamide	14.4	µg FluEQ/L
	GR	Glucocorticoid activity	dexamethasone	100	ng DexEQ/L
in vitro antibiotics	T	Bacterial growth inhibition (Tetracyclines)	oxytetracycline	250	ng OxyEQ/L
	Q	Bacterial growth inhibition (Quinolones)	flumequine	100	ng FlqEQ/L
	B+M	Bacterial growth inhibition (β-lactams and Macrolides)	penicillin G	50	ng PenEQ/L
	S	Bacterial growth inhibition (Sulfonamides)	sulfamethoxazole	100	ng SulEQ/L
	A	Bacterial growth inhibition (Aminoglycosides)	neomycin	500	ng NeoEQ/L

assay. A specific water sample induced an effect of 38% in this assay, equivalent to the effect of approximately 0.02 nM bioanalytical equivalents.

Effect-based trigger values

The identification of ecological risks from bioassay battery responses follows from the comparison of bioanalytical signals to previously determined thresholds, defined as effect-based trigger values (EBT), that should differentiate between acceptable and poor water quality. Since bioassays potentially respond to the mixture of all compounds present in a sample, effect-based trigger values are expressed as toxic or bioanalytical equivalents of concentrations of model compounds for the respective bioassay (Table 1).

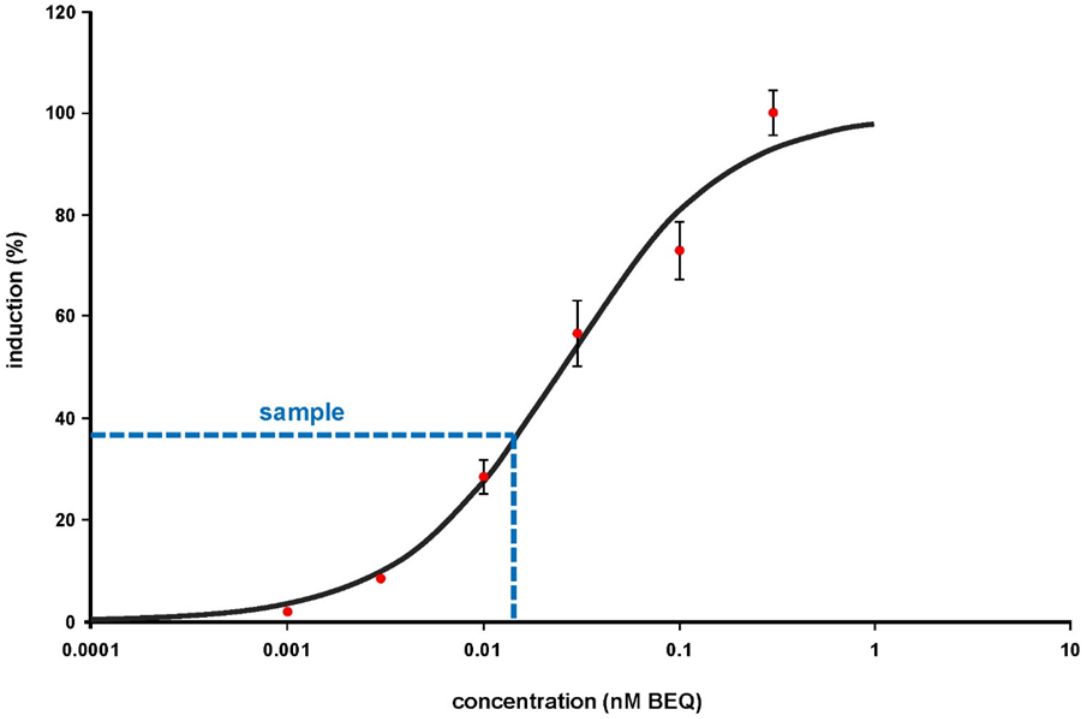


Figure 2. Dose response relationship for a reference compound in an in vitro bioassay. The blue lines show that a specific water sample induced an effect of 38%, representing approximately 0.02 nM bioanalytical equivalents.

Ranking of contaminated sites based on effect-based risk assessment

Once the toxic potency of a sample in a bioassay is expressed as toxic equivalent concentrations or bioanalytical equivalent concentrations, this response can be compared to the effect-based trigger value for that assay, thus determining whether or not there is a potential ecological risk from contaminants in the investigated water sample. The ecotoxicity profiles of the surface water samples generated by a bioassay battery allow for calculation and ranking of a cumulative ecological risk for the selected locations. In the example given in Figure 3, water samples of six locations were subjected to the SIMONI bioassay battery of Van der Oost et al. (2017), consisting of 17 *in situ*, *in vivo* and *in vitro* bioassays. Per site and per bioassay the response is compared to the corresponding effect-based trigger value and classified as 'no response' (green), 'response below the effect-based trigger value' (yellow) or 'response above the effect-based trigger value' (orange). Next, the cumulative ecological risk per location is calculated.

The resulting integrated ecological risk score allows ranking of the selected sites based on the presence of ecotoxicological risks rather than on the presence of a limited number of target compounds. This in turn permits water authorities to invest money where it matters most: identification of compounds causing adverse effects at locations with indicated ecotoxicological risks. Initially, the compounds causing the observed effect-based trigger value exceedance will not be known, however, this can subsequently be elucidated with targeted or non-target chemical analysis, which will only be necessary at locations with indicated ecological risks. A potential follow-up step could be to investigate the drivers of the observed effects by means of effect-directed analysis (see section on Effect-directed analysis).

	Location					
Bioassay	**HN2**	**WN3**	**WN6**	**DO3**	**PM2**	**BD3**
Daphnia in situ	4.25	0.50	2.50	0.50	0.00	0.00
Microtox	0.00	0.00	0.04	0.02	0.00	0.00
Algatox	0.00	0.00	0.00	0.01	0.00	0.00
Daphniatox	0.00	0.00	0.00	0.02	0.04	0.00
cytotox CALUX polar	2.68	0.18	0.60	0.60	0.79	0.18
cytotox CALUX non polar	0.01	0.00	0.00	0.00	0.00	0.00
ER CALUX	1.43	5.71	0.63	0.71	0.32	0.00
a-AR CALUX	10.00	2.05	1.17	1.31	0.54	0.74
GR CALUX	0.00	0.00	0.00	0.49	0.00	0.00
Antibiotics	0.04	0.14	0.01	0.62	0.02	0.00
DR CALUX	0.65	0.83	0.41	0.26	0.12	0.15
PAH CALUX	0.24	0.48	0.26	0.15	0.07	0.06
PPAR CALUX	0.00	0.00	0.00	0.00	0.00	0.00
Nrf2 CALUX	1.52	0.56	0.51	0.29	0.53	0.30
PRX CALUX	1.25	1.88	1.16	0.61	0.64	0.81
p53 CALUX -S9	0.06	0.03	0.03	0.01	0.02	0.00
p53 CALUX +S9	0.00	0.00	0.00	0.00	0.00	0.00
Integrated risk score	**2.63**	**1.29**	**0.98**	**0.62**	**0.31**	**0.22**

Figure 3. Heat map showing the response of 17 in situ, in vivo and in vitro bioassays to six surface water samples. The integrated risk score (SIMONI Risk Indication; Van der Oost et al., 2017) is classified as 'low risk' (green), 'potential risk' (orange) or 'risk' (red).

References

Escher, B. I., Aït-Aïssa, S., Behnisch, P. A., Brack, W., Brion, F., Brouwer, A., et al. (2018). Effect-based trigger values for in vitro and in vivo bioassays performed on surface water extracts supporting the environmental quality standards (EQS) of the European Water Framework Directive. Science of the Total Environment 628-629, 748-765.

Van der Oost, R., Sileno, G., Suarez-Munoz, M., Nguyen, M.T., Besselink, H., Brouwer, A. (2017). SIMONI (Smart Integrated Monitoring) as a novel bioanalytical strategy for water quality assessment: part I - Model design and effect-based trigger values. Environmental Toxicology and Chemistry 36, 2385-2399.

Additional reading

Altenburger, R., Ait-Aissa, S., Antczak, P., Backhaus, T., Barceló, D., Seiler, T.-B., et al. (2015). Future water quality monitoring - Adapting tools to deal with mixtures of pollutants in water resource management. Science of the Total Environment 512-513, 540-551.

Escher, B.I., Leusch, F.D.L. (2012). *Bioanalytical Tools in Water Quality Assessment*. IWA publishing, London (UK).

Hamers, T., Legradi, J., Zwart, N., Smedes, F., De Weert, J., Van den Brandhof, E-J., Van de Meent, D., De Zwart, D. (2018). Time-Integrative Passive sampling combined with TOxicity Profiling (TIPTOP): an effect-based strategy for cost-effective chemical water quality assessment. Environmental Toxicology and Pharmacology 64, 48-59.

6.4.4. Question 1

List 5 advantages of an effect-based approach over a compound based approach for water quality assessment.

6.4.4. Question 2

Motivate the necessity of employing a bioassay battery in effect based monitoring approaches.

6.4.4. Question 3

Explain how bioassay responses are expressed in terms of toxicity equivalents of reference compounds? (you may wish to draw a figure).

6.4.4. Question 4

Translate the outcome of a bioassay battery into a ranking of contaminated sites based on ecotoxicological risk. (you may wish to draw a figure).

6.4.5. Biomonitoring: in situ bioassays and contaminant concentrations in organisms

Author: Michiel Kraak

Reviewers: Ad Ragas, Suzanne Stuijfzand, Lieven Bervoets

Learning objectives:

You should be able to

- name tools specifically designed for ecological risk assessment in the field.
- define biomonitoring and to describe biomonitoring procedures.
- list the characteristics of suitable biomonitoring organisms.
- list the most commonly used biomonitoring organisms per environmental compartment.
- argue the advantages and disadvantages of *in situ* bioassays.
- argue the advantages and disadvantages of measuring contaminant concentrations in organisms.

Key words: Biomonitoring, test organisms, *in situ* bioassays, contaminant concentrations in organisms, environmental quality

Introduction

Several approaches and tools are available for diagnostic risk assessment. Tools specially developed for field assessments include the TRIAD approach (see section on TRIAD approach), in situ bioassays and biomonitoring. In ecotoxicology, biomonitoring is defined as the use of living organisms for the *in situ* assessment of environmental quality. Passive biomonitoring and active biomonitoring are distinguished. For passive biomonitoring, organisms are collected at the site of interest and their condition is assessed or the concentrations of specific target compounds in their tissues are analysed, or both. By comparing individuals from reference and contaminated sites an indication of the impact on local biota at the site of interest is obtained. For active biomonitoring, organisms are collected from reference sites and exposed in cages or artificial substrates at the study sites. Ideally, reference organisms are simultaneously exposed at the site of origin to control for potential effects of the experimental set-up on the test organisms. As an alternative to field collected animals, laboratory cultured organisms may be employed. After exposure at the study sites for a certain period of time, the organisms are recollected and either their condition is analysed (*in situ* bioassay) or the concentrations of specific target compounds are measured in the organisms, or both.

The results of biomonitoring studies may be used for management decisions, e.g. when accumulation of contaminants has been demonstrated in the field and especially when the sources of the pollution have been identified. However, the use of biomonitoring studies in environmental management has not been captured in formal protocols or guidelines like those of the Water Framework Directive (WFD) or - to a lesser extent - the TRIAD approach and effect-based quality assessments. Biomonitoring studies are typically applied on an case-by-case basis and their application therefore strongly depends on the expertise and resources available for the assessment. The text below explains and discusses the most important aspects of biomonitoring techniques used in diagnostic risk assessment.

Selection of biomonitoring test organisms

The selection of adequate organisms for biomonitoring partly follows the selection of test organisms for toxicity tests (see section on the Selection of test organisms). Suitable biomonitoring organisms:

- Are sedentary, since sedentary organisms may adapt more easily to the *in situ* experimental setup than more mobile organisms, for which caging may be an additional stress factor. Moreover, for sedentary organisms the relationship between the

accumulated compounds and the environmental quality at the exposure site is straightforward, although this is more relevant to passive than to active biomonitoring.

- Are representative for the community of interest and native to the study sites, since this will ensure that the biomonitoring organisms tolerate local conditions other than contamination, preventing that stressors other than poor environmental conditions may affect their performance. Obviously, it is also undesirable to introduce exotic species into new environments.
- Are long living, at least substantially longer than the exposure duration and preferably large enough to obtain sufficient material for chemical analysis.
- Are easy to handle.
- Respond to a gradient of environmental quality, if the purpose of the biomonitoring study is to analyse the condition of the organisms after recollection (*in situ* bioassay).
- Accumulate contaminants without being killed, if the purpose of the biomonitoring study is to measure contaminant concentrations in the organisms after recollection.
- Are large enough to obtain sufficient biomass for the analysis of the target compounds above the limits of detection, if the purpose of the biomonitoring study is to measure contaminant concentrations in the organisms after recollection.

Based on the above listed criteria, in the marine environment mussels belonging to the genus *Mytilus* are predominantly selected. The genus *Mytilus* has the additional advantage of a global distribution, although represented by different species. This facilitates the comparison of contaminant concentrations in the organisms all around the globe. Lugworms have occasionally also been used for biomonitoring in marine systems. For freshwater, the cladoceran *Daphnia magna* is most frequently employed, although occasionally other species are selected, including mayflies, snails, worms, amphipods, isopods, caddisflies and fish. Given the positive experience with marine mussels, freshwater bivalves are also employed as biomonitoring organisms. Sometimes primary producers have been used, mainly periphyton. Due to the complexity of the sediment and soil compartments, few attempts have been made to expose organisms *in situ*, mainly restricted to chironomids on sediment.

In situ exposure devices

An obvious requirement of the *in situ* exposure devices is that the test organisms do not suffer from (sub)lethal effects of the experimental setup. If the organisms are large enough, cages may be used, like for freshwater and marine mussels. For daphnids, a simple glass jar with a permeable lid suffices. For riverine insects, the device should allow the natural flow of the stream to pass, but meanwhile prevent the organisms from escaping. In the device shown in Figure 1a, caddisfly larvae containing tubes are connected to floating tubes, maintaining the larvae at a constant depth of 65 cm. In the tubes, the caddisfly larvae are able to settle and build nets on artificial substrate, a plastic doormat with bristles standing out.

An elegant device for *in situ* colonization of periphyton was developed by Blanck (1985)(Figure 1b). Sand-blasted glass discs (1.5 cm^2 surface) are used as artificial substratum for algal attachment. Substrata are placed vertically in the water, parallel to the current, by means of polyethylene racks, each rack supporting a total of 170 discs. After the colonization period, the periphyton containing glass discs can be harvested, offering the unique possibility to perform laboratory or field experiments with entire algal and microbial communities, replicated 170 times.

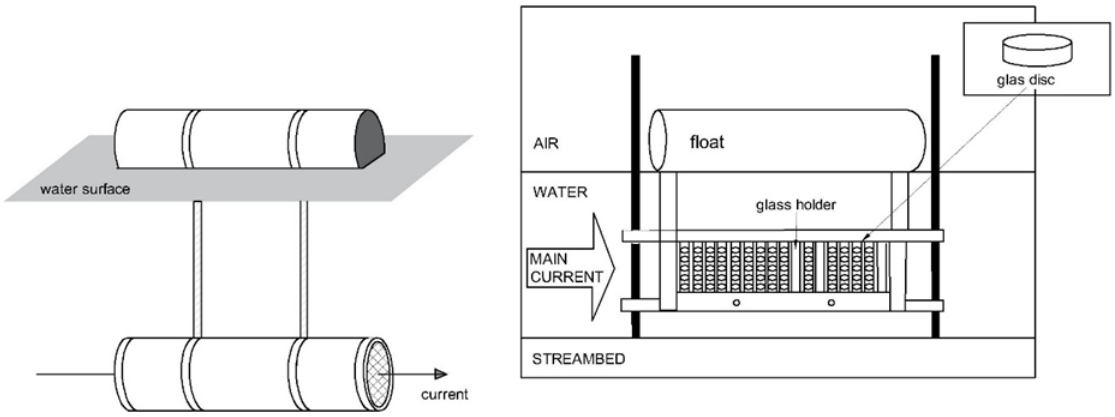


Figure 1. Left: Experimental set-up for in situ exposure of caddisfly larvae according to Stuijfzand et al. (1999, derived from Vuori (1995)). Right:

Experimental set-up for in situ colonization of periphyton according to Ivorra et al. (1999), derived from Blanck (1985). Drawn by Wilma IJzerman.

In situ bioassays

After exposure at the study sites for a certain period of time, the organisms are recollected and their condition can be analysed (Figure 2). The endpoint is mostly survival, especially in routine monitoring programs. If the *in situ* exposure lasts long enough, also effects on species specific sublethal endpoints can be assessed. For daphnids and snails, this is reproduction and for isopods growth. For aquatic insects (mayflies, caddisflies, damselflies, chironomids), emergence has been assessed as a sensitive ecological relevant endpoint (Barmentlo et al., 2018).

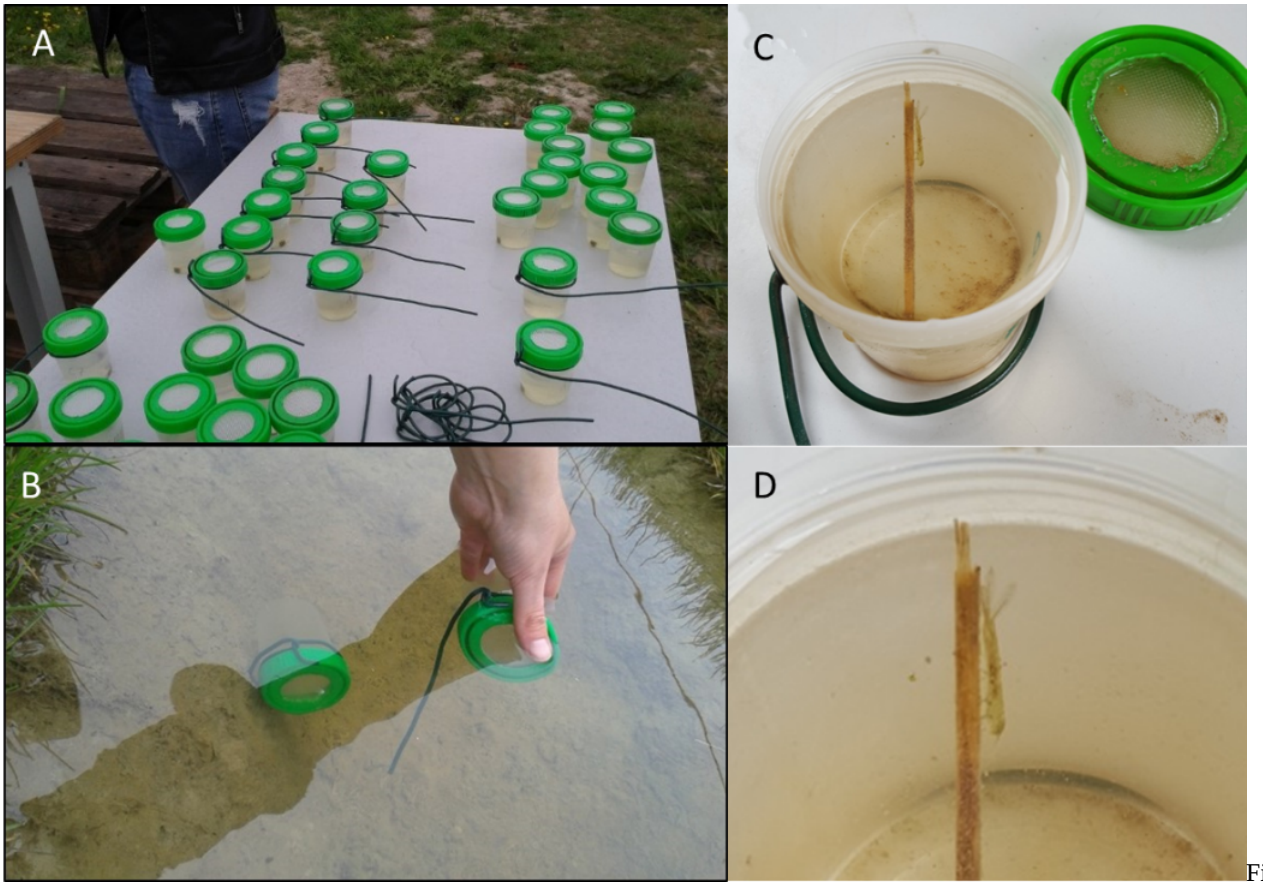


Figure 2. In situ exposure experiment. (A). Preparing damselfly containing jars. (B) Exposure of the in situ jars in ditches. (C). Retrieved jar containing a single damselfly larva. (D). Close up of the damselfly larva ready for inspection. Photos by Henrik Barmentlo.

In situ bioassays come closest to the actual field situation. Organisms are directly exposed at the site of interest and respond to all joint stressors present. Yet, this is also the limitation of the approach. If organisms do respond it remains unknown what causes the observed adverse effects. This could be (the combination of) any natural or anthropogenic physical or chemical stress factor. *In situ* bioassays can therefore be best combined with laboratory bioassays (see section on Bioassays) and the analysis of physico-chemical parameters, conform the TRIAD approach (see section on TRIAD approach). If the adverse effects are also observed in the bioassays under controlled laboratory conditions, then poor water quality is most likely the cause. The water sample may then be subjected to suspected target analysis, non-target analysis or effect directed analysis (EDA). If adverse effects are observed *in situ* but not in the laboratory, then the presence of hazardous compounds is most likely not the cause. Instead, the effects may be attributable to e.g. low pH, low oxygen concentrations, high temperatures etc, which may be verified by physico-chemical analysis in the field.

Online biomonitoring

A specific application of *in situ* bioassays are the online systems for continuous water quality monitoring. In these systems, behaviour is generally the endpoint (see section on Endpoints). Organisms are exposed in a laboratory setting *in situ* (on shore or

on a boat) in an experimental device to a continuous flow of surface water. If the water quality changes, the organisms respond by changing their behaviour. Above a certain threshold an alarm may go off and, for instance, the intake of surface water for drinking water preparation can be temporarily stopped.

Contaminant concentrations in organisms

As an addition or as an alternative to analysing the condition of the exposed biomonitoring organisms upon retrieval, contaminant concentrations in organisms can be analysed. This has several advantages over chemical analysis of environmental samples: biomonitoring organisms may be exposed for days to weeks at the site of interest, providing time integrated measurements of contaminant concentrations, in contrast to the chemical analysis of grab samples. This way, biomonitoring organisms actually serve as 'biological passive samplers' (see to section on Experimental methods of assessing available concentrations of organic chemicals). Another advantage of measuring contaminant concentrations in organisms is that they only take up the bioavailable (fraction of) substances, ecologically very relevant information, that remains unknown if chemical analysis is performed on water, sediment, soil or air samples. Yet, elevated concentrations in organisms do not necessarily imply toxic effects, and therefore these measurements are best complemented with determining their condition, as described above. Moreover, analysing contaminants in organisms may be more expensive than measurements of environmental samples, due to a more complex sample preparation. Weighing the advantages and disadvantages, the explicit strength of biomonitoring programs is that they provide insight into the spatial and temporal variation in bioavailable contaminant concentrations. In Figure 3 two examples are given. The left panel shows the concentrations of PCBs in zebra mussels at different sampling sites in Flanders, Belgium (Bervoets et al., 2004). The right panel shows the rapid (within 2 wk) Cd accumulation and depuration in biofilms translocated from a reference to a polluted site and from a polluted to a reference site, respectively (Ivorra et al., 1999).

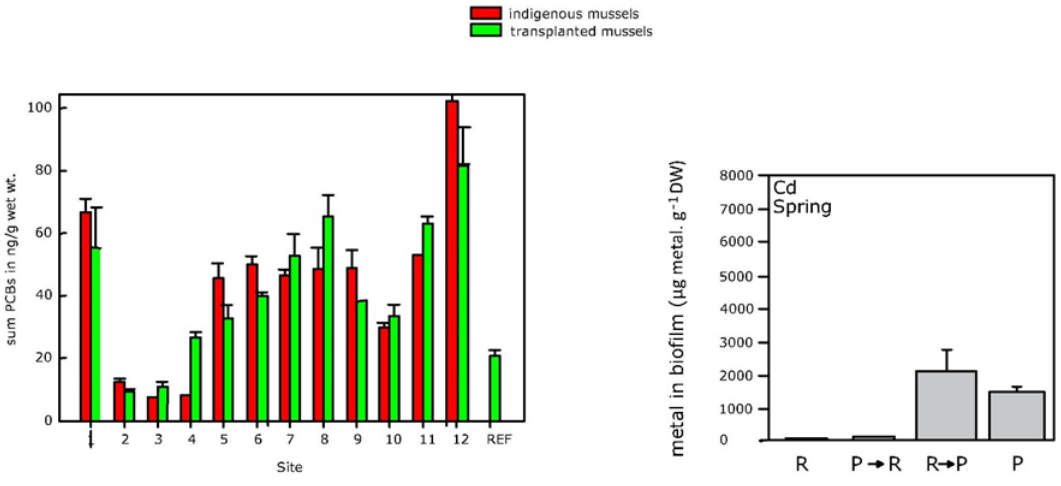


Figure 3. Left panel: Mean concentration of PCBs in 25 pooled zebra mussels at different sampling sites in Flanders, Belgium. Comparison between indigenous (black bars) and transplanted mussels (grey bars), from Bervoets et al. (2004). Right panel: Cd concentrations in local and translocated biofilms (R: Reference site; P: polluted site) from Ivorra et al. (1999). Drawn by Wilma IJzerman.

References

Barmentlo, S.H., Parmentier, E.M., De Snoo, G.R., Vijver, M.G. (2018). Thiacloprid-induced toxicity influenced by nutrients: evidence from *in situ* bioassays in experimental ditches. Environmental Toxicology and Chemistry 37, 1907-1915.

Bervoets, L., Voets, J., Chu, S.G., Covaci, A., Schepens, P., Blust, R. (2004). Comparison of accumulation of micropollutants between indigenous and transplanted zebra mussels (*Dreissena polymorpha*). Environmental Toxicology and Chemistry 23, 1973-1983.

Blanck, H. (1985). A simple, community level, ecotoxicological test systemusing samples of periphyton. Hydrobiologia 124, 251-261.

Ivorra, N., Hettelaar, J., Tubbing, G.M.J., Kraak, M.H.S., Sabater, S., Admiraal, W. (1999). Translocation of microbenthic algal assemblages used for *in situ* analysis of metal pollution in rivers. Archives of Environmental Contamination and Toxicology 37, 19-28.

Stuijfzand, S.C., Engels, S., Van Ammelrooy, E., Jonker, M. (1999). Caddisflies (Trichoptera: Hydropsychidae) used for evaluating water quality of large European rivers. Archives of Environmental Contamination and Toxicology 36, 186-192.

Vuori, K.M. (1995). Species- and population-specific responses of translocated hydropsychid larvae (Trichoptera, Hydropsychidae) to runoff from acid sulphate soils in the River Kyronjoki, western Finland. Freshwater Biology 33, 305-318.

6.4.5. Question 1

Define biomonitoring.

6.4.5. Question 2

Explain the difference between passive and active biomonitoring.

6.4.5. Question 3

What can be measured in biomonitoring organisms after (re)collection?

6.4.5. Question 4

Explain the difference between passive and active biomonitoring.

6.4.5. Question 5

List the characteristics of suitable biomonitoring organisms.

6.4.5. Question 6

Name the advantage and disadvantage of *in situ* bioassays.

6.4.5. Question 7

Name the advantages and disadvantages of measuring contaminant concentrations in organisms.

6.4.6. TRIAD approach for site-specific ecological risk assessment

Author: Michiel Rutgers

Reviewers: Kees van Gestel, Michiel Kraak, Ad Ragas

Learning goals:

You should be able

- to describe the principles of the TRIAD approach
- to explain the importance of weight of evidence in risk assessment
- to use the results for an assessment by applying the TRIAD approach

Keywords: Triad, site-specific ecological risk assessment, weight of evidence

Like the other diagnostic tools described in the previous sections (see sections on Effect-based monitoring In vivo bioassays and In vitro bioassays, Effect-directed analysis, and Effect-based water quality assessment and Biomonitoring), the TRIAD approach is a tool for site-specific ecological risk assessment of contaminated sites (Jensen et al., 2006; Rutgers and Jensen, 2011). Yet, it differs from the previous approaches by combining and integrating different techniques through a 'weight of evidence' approach. To this purpose, the TRIAD combines information on contaminant concentrations (environmental chemistry), the toxicity of the mixture of chemicals present at the site ((eco)toxicology), and observations of ecological effects (ecology) (Figure 1).

The mere presence of contaminants is just an indication of potential ecological effects to occur. Additional data can help to better assess the ecological risks. For instance, information on actual toxicity of the contaminated site can be obtained from the exposure of test organisms to (extracts of) environmental samples (bioassays), while information on ecological effects can be obtained from an inventory of the community composition at the specific site. When these disciplines tend to converge to corresponding levels of ecological effects, a weight of evidence is established, making it possible to finalize the assessment and to support a decision for contaminated site management.

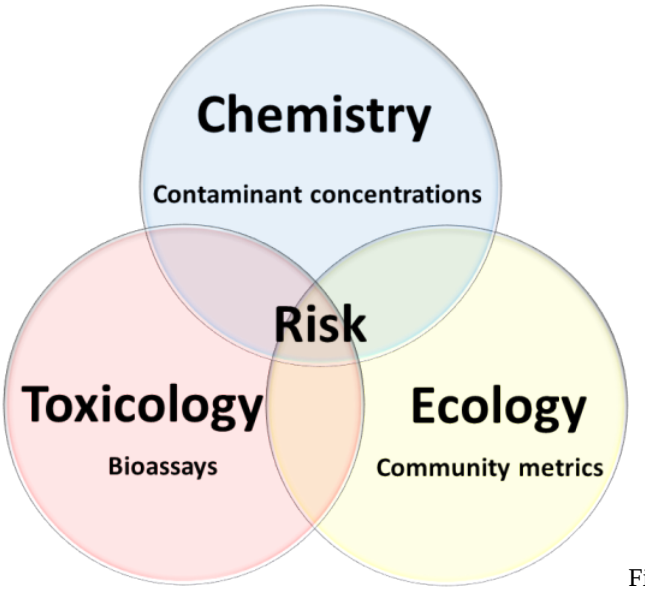


Figure 1: The TRIAD approach integrating information on contaminant concentrations (environmental chemistry), bioassays ((eco)toxicology) and ecological field inventories (ecology) into a weight of evidence for site-specific ecological risk assessment (adapted from Chapman, 1988).

The TRIAD approach thus combines the information obtained from three lines of evidence (LoE):

1. *LoE Chemistry*: risk information obtained from the measured contaminant concentrations and information on their fate in the ecosystem and how they can evoke ecotoxicological effects. This can include exposure modelling and bioavailability considerations.
2. *LoE Toxicity*: risk information obtained from (eco)toxicity experiments exposing test organisms to (extracted) samples of the site. These bioassays can be performed on site or in the laboratory, under controlled conditions.
3. *LoE Ecology*: risk information obtained from the observation of actual effects in the field. This is deduced from data of ecological field surveys, most often at the community level. This information may include data on the composition of soil communities or other community metrics and on ecosystem functioning.

The three lines of evidence form a weight of evidence when they are converging, meaning that when the independent lines of evidence are indicating a comparable risk level, there is sufficient evidence for providing advice to decision makers about the ecological risk at a contaminated site. When there is no convergence in risk information obtained from the three lines of evidence, uncertainty is large. Further investigations are then required to provide a unambiguous advice.

Table 1. Basic data for site-specific environmental risk assessment (SS-ERA) sorted per line of evidence (LoE). Data and methods are described in Van der Waarde et al. (2001) and Rutgers et al. (2001).

	Basic data	unit	Site A	Site B	Ref
chemistry	pH (KCl)		2.8	3.1	2.9
	SOM	%	2.4	2.1	3.2
	lutum	%	0.5	0.5	0.76
	lead concentration	mg/kg	290	880	11
	antimoon concentration	mg/kg	3	3.1	3
	arsenicum concentration	mg/kg	2.5	2.9	2
	zinc concentration	mg/kg	7.6	7.2	5.9
	toxic pressure (Pb,Sb,As, Zn)	msPAF	0.49	0.71	0
Toxicology	Microtox assay	% inhibition	95	95	0
	Lettuce germination	% inhibition	55	10	0
	Lettuce growth	% inhibition	26	60	0
Ecology	nematodes abundance	n/ 100 g soil	420	130	410
	nematodes Maturity Index 2-5	species composition	2.38	3.41	2.73
	Bait lamina	% activity	60	32	90

Tests and abbreviations used in the table:

- Toxic Pressure metals (sum TP metals). The toxic pressure of the mixture of metals in the sample, calculated as the potentially affected fraction in a Species Sensitivity Distribution with NOEC values (see Section on SSDs) and a simple rule for mixture toxicity (response addition; Section on Mixture toxicity).
- Microtox. A bioassay with the luminescent bacterium *Allivibrio fischeri*, formerly known as *Vibrio fischeri*. Luminescence is reduced when toxicity is high.
- Lettuce Growth and Lettuce Germination. A bioassay with the growth performance and the germination percentage of lettuce (seeds).
- Bait Lamina. The bait-lamina test consists of vertically inserting 16-hole-bearing plastic strips filled with a plant material preparation into the soil. This gives an indication of the feeding activity of soil animals.
- Nematodes abundance and Nematodes Maturity Index 2-5. The biomass and the Maturity Index (MI) of the nematode community in soil samples provide information about soil health (Van der Waarde et al. 2001).

The results of a site-specific ecological risk assessment (SS-ERA) applying the TRIAD approach are first organized basic tables for each sample and line of evidence separately. Table 1 shows an example. This table also collects supporting data, such as soil pH and organic matter content. Subsequently, these basic data are processed into ecological risk values by applying a risk scale running from zero (no effects) to one (maximum effect). An example of a metric used is the multi-substance Potentially Affected Fraction of species from the mixture of contaminants (see Section on SSDs). These risk values are then collected in a TRIAD table (Table 2), for each endpoint separately, integrated per line of evidence individually, and finally integrated over the three lines of evidence. Also the level of agreement between the three lines of evidence is given a score. Weighting values are applied, e.g. equal weights for all ecological endpoints (depending on number of methods and endpoints), and equal weights for each line of evidence (33%). When differential weights are preferred, for instance when some data are judged as unreliable, or some endpoints are considered more important than others, the respective weight factors and the arguments to apply them must be provided in the same table and accompanying text.

Table 2. Soil Quality TRIAD table demonstrating scaled risk values for two contaminated sites (A, B) and a Reference site (based on real data, only for illustration purposes). Risk values are collected per endpoint, grouped according to respective Lines of Evidence (LoE), and finally integrated into a TRIAD value for risks. The deviation indicates a level of agreement between LoE (default threshold 0.4). For site B, a Weight of Evidence (WoE) is demonstrated (D<0.4) making decision support feasible. By default equal weights can be used throughout. Differential weights should be indicated in the table and described in the accompanying text.

Triad LoE	Parameter	Site A	Site B	Ref
Chemistry	Sum TP metals	0.49	0.71	0.00
	integration	**0.49**	**0.77**	**0.00**
Toxicology	Microtox	0.95	0.95	0.00
	Lettuce germination	0.55	0.09	0.00
	Lettuce growth	0.26	0.60	0.00
	integration	**0.59**	**0.55**	**0.00**
Ecology	Nematodes abundance	0.00	0.68	0.00
	Nematodes Maturity Index	0.00	0.30	0.00
	Bait lamina	0.30	0.64	0.00
	integration	**0.10**	**0.54**	**0.00**
	chemical risk	0.49	0.71	0.00
	ecotoxicity risk	0.59	0.55	0.00
	ecology risk	0.10	0.54	0.00
	final score for risk	***0.42***	***0.61***	***0.00***
	deviation	0.44	0.17	0.00

References

ISO (2017). ISO 19204: Soil quality -- Procedure for site-specific ecological risk assessment of soil contamination (soil quality TRIAD approach). International Standardization Organization, Geneva. https://www.iso.org/standard/63989.html.

Jensen, J., Mesman, M. (Eds.) (2006). LIBERATION, Ecological risk assessment of contaminated land, decision support for site specific investigations. ISBN 90-6960-138-9, Report 711701047, RIVM, Bilthoven, The Netherlands.

Rutgers, M., Bogte, J.J., Dirven-Van Breemen, E.M., Schouten, A.J. (2001) Locatiespecifieke ecologische risicobeoordeling - praktijkonderzoek met een Triade-benadering. RIVM-rapport 711701026, Bilthoven.

Rutgers, M., Jensen, J. (2011). Site-specific ecological risk assessment. Chapter 15, in: F.A. Swartjes (Ed.), Dealing with Contaminated Sites - from Theory towards Practical Application, Springer, Dordrecht. pp. 693-720.

Van der Waarde, J.J., Derksen, J.G.M, Peekel, A.F., Keidel, H., Bloem, J., Siepel, H. (2001) Risicobeoordeling van bodemverontreiniging met behulp van een triade benadering met chemische analyses, bioassays en biologische veldinventarisaties. Eindrapportage NOBIS 98-1-28, Gouda.

6.4.6. Question 1

A sediment sample was analyzed for Priority Hazardous Substances (PHSs), but none were detected. Yet, this sediment sample caused high mortality in laboratory bioassays with three sediment inhabiting species. At the site where the sample was taken biodiversity was very low. Explain these observations.

6.4.6. Question 2

In a sediment sample, the total concentration of Priority Hazardous Substances (PHSs) was shown to be very high. Yet, this sediment sample caused no mortality in laboratory bioassays with three sediment inhabiting species. Moreover, at the sample site in the field species rich communities were observed. Explain these observations.

6.4.6. Question 3

What is the added value of using bioassays and field observations over chemical analysis when assessing the potential risk of a contaminated site?

6.4.6. Question 4

What is the added value of performing an assessment along three independent Lines of Evidence (LoE)?

6.4.7. Eco-epidemiology

Authors: Leo Posthuma, Dick de Zwart

Reviewers: Allan Burton, Ad Ragas

Learning objectives:

You should be able to:

- explain that and how effects of chemicals and their mixtures can be demonstrated in monitoring data sets;
- explain that effects can be characterized with various impact metrics;
- formulate whether and how the choice of impact sensitivity metric is relevant for the sensitivity and outcomes of a diagnostic assessment;
- explain how ecological and ecotoxicological analysis methods relate;
- explain how eco-epidemiological analyses are helpful in validating ecotoxicological models utilized in ecotoxicological risk assessment and management.

Keywords: eco-epidemiology, mixture pollution, diagnosis, impact magnitude, probable causes, validation

Introduction

Approaches for environmental protection, assessment and management differ between 'classical' stressors (such as excess nutrients and pH) and chemical pollution. For the 'classical' environmental stress factors, ecologists use monitoring data to develop concepts and methods to prevent and reduce impacts. Although there are some clear-cut examples of chemical pollution impacts [e.g., the decline in vulture populations in South East Asia due to diclofenac (Oaks et al. 2004), and the suit of examples in the book 'Silent Spring' (Carson 1962)], ecotoxicologists commonly have assessed the stress from chemical pollution by evaluating exposures vis a vis laboratory toxicity data. Current pollution often consists of complex mixtures of chemicals, with highly variable patterns in space and time. This poses problems when one wants to evaluate whether observed impacts in ecosystems can be attributed to

chemicals or their mixtures. Eco-epidemiological methods have been established to discern such pollution stress. These methods provide the diagnostic tools to identify the impact magnitude and key chemicals that cause impacts in ecosystems. The use of these methods is further relevant for validating the laboratory-based risk assessment approaches developed by ecotoxicology.

The origins of eco-epidemiology

Risk assessments of chemicals provide insights in expected exposures and impacts, commonly for separate chemicals. These are predictive outcomes with a high relevance for decision making on environmental protection and management. The validation of those risk assessments is key to avoid wrong protection and management decisions, but it is complex. It consists of comparing predicted risk levels to observed effects. This begs the question on how to discern effects of chemical pollution in the field. This question can be answered based on the principles of ecological bio-assessments combined with those of human epidemiology. A bio-assessment is a study of stressors and ecosystem attributes, made to delineate causes of impacts via (often) statistical associations between biotic responses and particular stressors. Epidemiology is defined as the study of the distribution and causation of health and disease conditions in specified populations. Applied epidemiology serves as a scientific basis to help counteracting the spreading of human health problems. Dr. John Snow is often referred to as the 'father of epidemiology'. Based on observations on the incidence, locations and timings of the 1854 cholera outbreak in London, he attributed the disease to contaminated water taken from the Broad Street pump well, counteracting the prevailing idea that the disease was caused by transmission via air. His proposals to control the disease were effective. Likewise, eco-epidemiology - in its ecotoxicological context - has been defined as the study of the distribution and causation of impacts of multiple stressor exposures in ecosystems. In its applied form, it supports the reduction of ecological impacts of chemical pollution. Human-health eco-epidemiology is concerned with environment-mediated disease.

The first literature mention of eco-epidemiological analyses on chemical pollution stems from 1984 (Bro-Rasmussen and Løkke 1984). Those authors described eco-epidemiology as a discipline necessary to validate the risk assessment models and approaches of ecotoxicology. In its initial years, progress in eco-epidemiological research was slow due to practical constraints such as a lack of monitoring data, computational capacity and epidemiological techniques.

Current eco-epidemiology

Current eco-epidemiological studies in ecotoxicology aim to diagnose the impacts of chemical pollution in ecosystems, and utilize a combination of approaches in order to diagnose the role of chemical mixtures in causing ecological impacts in the field. The combination of approaches consists of:

1. Collection of monitoring data on abiotic characteristics and the occurrence and/or abundance of biotic species, for the environmental compartment under study;

2. If needed: data optimization, usually to align abiotic and biotic monitoring data, including the chemicals;

3. Statistical analysis of the data set using eco-epidemiological techniques to delineate impacts and probable causes, according to the approaches followed in 'classical' ecological bio-assessments;

4. Interpretation and use of the outcomes for either validation of ecotoxicological models and approaches, or for control of the impacts sensu Dr. Snow.

Key examples of chemical effects in nature

Although impacts of chemicals in the environment were known before 1962, Rachel Carson's book Silent Spring (see Section on the history of Environmental toxicology) can be seen as early and comprehensive eco-epidemiological study that synthesized the available information of impacts of chemicals in ecosystems. She considered effects of chemicals a novel force in natural selection when she wrote: "If Darwin were alive today the insect world would delight and astound him with its impressive verification of his theories of survival of the fittest. Under the stress of intensive chemical spraying the weaker members of the insect populations are being weeded out."

Clear examples of chemical impacts on species are still reported. Amongst the best-known examples is a study on vultures. The population of Indian vultures declined more than 95% due to diclofenac exposure which was used intensively as a veterinary drug (Oaks et al. 2004). The analysis of chemical impacts in nature becomes however more complex over time. The diversity of

chemicals produced and used has vastly increased, and environmental samples contain thousands of chemicals at often low concentrations. Hence, contemporary eco-epidemiology is complex. Nonetheless, various studies demonstrated that contemporary mixture exposures affect species assemblages. Starting from large-scale monitoring data and following the four steps mentioned above, De Zwart et al. (2006) were able to show that effects on fish species assemblages could be attributed to both habitat characteristics and chemical mixtures. Kapo and Burton Jr (2006) showed the impacts of multiple stressors and chemical mixtures in aquatic species assemblages with similar types of data, but slightly different techniques. Eco-epidemiological studies of the effects of chemicals and their mixtures currently represent different geographies, species groups, stressors and chemicals/mixtures that are considered. The potential utility eco-epidemiological studies was reviewed by Posthuma et al. (2016). The review showed that mixture impacts occur, and that they can be separated from natural variability and multiple-stressor impacts. That means that water managers can develop management plans to counteract stressor impacts. Thereby, the study outcomes are used to prioritize management to sites that are most affected, and to chemicals that contribute most to those effects. Based on sophisticated statistical analyses, Berger et al. (2016) suggested chemicals can induce effects in the environment at concentrations much lower than expected based on laboratory experiments. Schäfer et al. (2016) argued that eco-epidemiological studies that cover both mixtures and other stressors are essential for environmental quality assessment and management. In practice, however, the analysis of the potential impacts of chemical mixtures is often still separate from the analysis of impacts of other stressors.

Steps in eco-epidemiological analysis

Various regulations require collection of monitoring data, followed by bio-assessment, such as the EU Water Framework Directive (see section on the Water Framework Directive). Therefore, monitoring data sets are increasingly available. The data set is subsequently curated and/or optimized for the analyses. Data curation and management steps imply amongst others that taxonomic names of species are harmonized, and that metrics for abiotic and biotic variables represent the conditions for the same place and time as much as possible. Next, the data set is expanded with novel variables, e.g. a metric for the toxic pressure exerted by chemical mixtures. An example of such a metric is the multi-substance Potentially Affected Fraction of species (msPAF). This metric transfers measured or predicted concentrations into the Potentially Affected Fraction of species (PAF), the values of which are then aggregated for a total mixture (De Zwart and Posthuma 2005). This is crucial, as adding each chemical of interest as a separate variable implies an increasingly expanding number of required sampling sites to maintain statistical power to diagnose impacts and probable causation.

The interpretation of the outcomes of the statistical analyses of the data set is the final step. Here, it must be acknowledged that statistical association is not equal to causation, and that care must be taken to explain the findings as indicative for mixture effects. Depending on the context of the study, this may then trigger a refined assessment, or alignment with other methods to collect evidence, or a direct use in an environmental management program.

Eco-epidemiological methods

A very basic eco-epidemiological method is quantile regression. Whereas common regression methods explore the magnitude of the change of the mean of the response variable (e.g., biodiversity) in relation to a predictor variable (e.g., pollutant stress), the quantile regression looks at the tails of the distributions of the response variable. How this principle operates is illustrated in Figure 1. When a monitoring data set contains one stressor variable at different levels (i.e., a gradient of data), the observations typically take the shape of a common stressor-response relationship (see section on Concentration-effect relationships). If the monitoring sites are affected by an extra stressor, the maximum-performance under the first stressor cannot be reached, so that the area under the curve contains the XY-points for this situation. Further addition of stressor variables and levels fills this space under the curve. When the raw data plotted as XY show an 'empty area' lacking XY-points, e.g. in the upper right corner, it is likely that the stressor variable can be identified as a stressor that limits the response variable, for example: chemicals limit biodiversity. The quantile regression calculates an upper percentiles (e.g., the 95th percentile) of the Y-values in assigned subgroups of X-values ("bins"). Such a procedure yields a picture such as Figure 1.

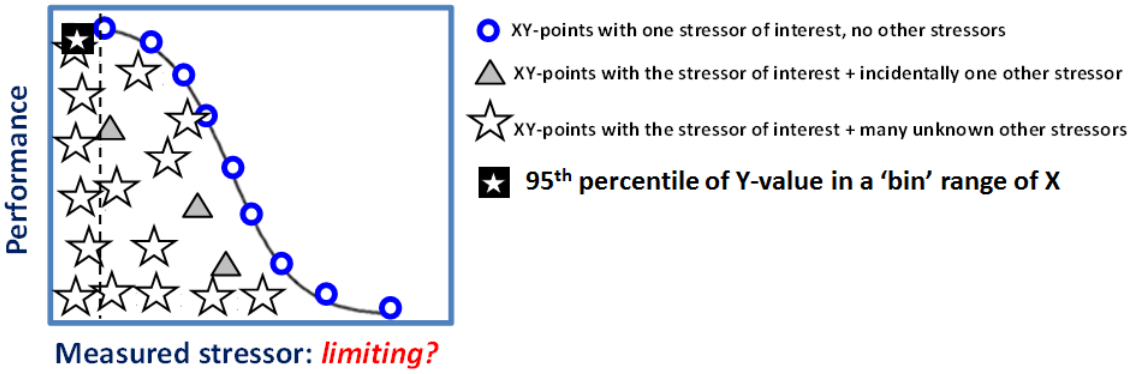


Figure 1. The principle of quantile regression in identification of a predictor variable (= stressor) that acts as a limiting factor to a response variable (= performance). It is common to derive e.g. the 95th percentile of the Y values in a 'bin' of X values to derive a stressor-impact curve. As illustration, the 95th percentile is marked only for the first bin of X values, with the blackened star.

More complex methods for analysis of (bio)monitoring data have been developed and applied. The methods are closely associated to those developed for, and utilized in, applied ecology. Well-known examples are 'species distribution models' (SDM), which are used to describe the abundance or presence of species as a function of multiple environmental variables. A well-known SDM is the bell-shaped curve relating species abundances to water pH: numbers of individuals of a species are commonly low at low and high pH, and the SDM is characterized as an optimum model for species abundance (Y) versus pH (X). Statistical models can also describe species abundance, presence or biodiversity, as a function of multiple stressors, for example via Generalized Linear Models. These have the general shape of:

$$Log(Abundance) = (a.\ pH + a'\ pH^2) + (b.\ OM + b'\ OM^2) + \ldots\ldots + e,$$

with a, a', b and b' being estimated from fitting the model to the data, whilst pH and OM are the abiotic stressor variables (acidity and Organic Matter, respectively); the quadratic terms are added to allow for optimum and minimum shaped relationships. When SSD models (see Section on Species Sensitivity Distribution) are used to predict the multi-substance Potentially Affected Fraction of species, the resulting mixture stress proxy can be analysed together with the other stressor variables. Data analyses from monitoring data from the United States and the Netherlands have, for example, shown that the abundance of >60% of the taxa is co-affected by mixtures of chemicals. An example study is provided by Posthuma et al. (2016).

Prospective mixture impact assessments

In addition to the retrospective analysis of monitoring data, in search of chemical impacts, recent studies also show examples of prospective studies of effects of mixtures. Different land uses imply different chemical use patterns, summarized as 'signatures'. That is, agricultural land use will yield intermittent emissions of crop-specific plant protection products, aligning with the growing season. Emissions from populated areas will show continuous emission of household chemicals and discontinuous emissions of chemicals in street run-off associated to heavy rain events. The application of emission, fate and ecotoxicity models showed that aquatic ecosystems are subject to the 'signatures', with associated predicted impact magnitudes (Holmes et al. 2018; Posthuma et al. 2018). Although such prospective assessments did not yet prove ecological impacts, they can assist in avoiding impacts by preventing the emission 'signatures' that are identified as potentially most hazardous.

The use of eco-epidemiological output

Eco-epidemiological analysis outputs serve two purposes, closely related to prospective and retrospective risk assessment of chemical pollution:

1. Validation of ecotoxicological models and approaches;

2. Derivation of control measures, to reduce impacts of diagnosed probable causes of impacts.

If needed, multiple lines of evidence can be combined, such as in the Triad approach (see section on TRIAD) or approaches that consider more than three lines of evidence (Chapman and Hollert, 2006). The higher the importance of a good diagnosis, the better

the user may rely on multiple lines of evidence.

First, the validation of ecotoxicological models and approaches is crucial, to avoid that important environmental protection, assessment and management activities rely on approaches that have limited relationship to field effects. Eco-epidemiological analyses have, for example, been used to validate the protective benchmarks used in the chemical-oriented environmental policies.

Second, the outcomes of an eco-epidemiological analysis can be used to control causes of impacts to ecosystems. Some studies have, for example, identified a statistical association between observed impacts (species expected but absent) and pollution of surface waters with mixtures of metals. Though local experts first doubted this association due to lack of industrial activities with metals in the area, they later found the association relevant given the presence of old spoil heaps from past mining activities. Metals appeared to leach into the surface waters at low rates, but the leached mixtures appeared to co-vary with species missing (De Zwart et al. 2006).

References

Berger, E., Haase, P., Oetken, M., Sundermann, A. (2016). Field data reveal low critical chemical concentrations for river benthic invertebrates. Science of The Total Environment 544, 864-873.

Bro-Rasmussen, F., Løkke, H. (1984). Ecoepidemiology - a casuistic discipline describing ecological disturbances and damages in relation to their specific causes; exemplified by chlorinated phenols and chlorophenoxy acids. Regulatory Toxicology and Pharmacology 4, 391-399.

Carson, R. (1962). Silent spring. Boston, Houghton Mifflin.

Chapman, P.M., Hollert, H. (2006). Should the sediment quality triad become a tetrad, a pentad, or possibly even a hexad? Journal of Soils and Sediments 6, 4-8.

De Zwart, D., Dyer, S.D., Posthuma, L., Hawkins, C.P. (2006). Predictive models attribute effects on fish assemblages to toxicity and habitat alteration. Ecological Applications 16, 1295-1310.

De Zwart, D., Dyer, S.D., Posthuma, L., Hawkins, C.P. (2006). Use of predictive models to attribute potential effects of mixture toxicity and habitat alteration on the biological condition of fish assemblages. Ecological Applications 16, 1295-1310.

De Zwart, D., Posthuma, L. (2005). Complex mixture toxicity for single and multiple species: Proposed methodologies. Environmental Toxicology and Chemistry 24,: 2665-2676.

Holmes, C.M., Brown, C.D., Hamer, M., Jones, R., Maltby, L., Posthuma, L., Silberhorn, E., Teeter, J.S., Warne, M.S.J., Weltje, L. (2018). Prospective aquatic risk assessment for chemical mixtures in agricultural landscapes. Environmental Toxicology and Chemistry 37, 674-689.

Kapo, K.E., Burton Jr, G.A. (2006). A geographic information systems-based, weights-of-evidence approach for diagnosing aquatic ecosystem impairment. Environmental Toxicology and Chemistry 25, 2237-2249.

Oaks, J.L., Gilbert, M., Virani, M.Z., Watson, R.T., Meteyer, C.U., Rideout, B.A., Shivaprasad, H.L., Ahmed, S., Chaudhry, M.J., Arshad, M., Mahmood, S., Ali, A., Khan, A.A. (2004). Diclofenac residues as the cause of vulture population decline in Pakistan. Nature 427(6975), 630-633.

Posthuma, L., Brown, C.D., de Zwart, D., Diamond, J., Dyer, S.D., Holmes, C.M., Marshall, S., Burton, G.A. (2018). Prospective mixture risk assessment and management prioritizations for river catchments with diverse land uses. Environmental Toxicology and Chemistry 37, 715-728.

Posthuma, L., De Zwart, D., Keijzers, R., Postma, J. (2016). Water systems analysis with the ecological key factor 'toxicity'. Part 2. Calibration. Toxic pressure and ecological effects on macrofauna in the Netherlands. Amersfoort, the Netherlands, STOWA.

Posthuma, L., Dyer, S.D., de Zwart, D., Kapo, K., Holmes, C.M., Burton Jr, G.A. (2016). Eco-epidemiology of aquatic ecosystems: Separating chemicals from multiple stressors. Science of The Total Environment 573, 1303-1319.

Posthuma, L., Suter, II, G.W., Traas, T.P. (Eds.) (2002). Species Sensitivity Distributions in Ecotoxicology. Boca Raton, FL, USA, Lewis Publishers.

Schäfer, R.B., Kühn, B., Malaj, E., König, A., Gergs, R. (2016). Contribution of organic toxicants to multiple stress in river ecosystems. Freshwater Biology 61, 2116-2128

6.4.7. Question 1

Which motivations do you know for executing eco-epidemiological analyses?

6.4.7. Question 2

Pollution in the 1950s and 1960s showed clear evidence for major impacts of chemicals in nature, whilst regulatory management actions taken since then have reduced clear recognition of chemical impacts in nature. Which ecotoxicological model has rejuvenated the development of eco-epidemiological methods?

6.4.7. Question 3

What is a simple approach, via which even raw-data plottings already show whether a stress factor (such as chemical mixture exposure) is limiting for ecology

---

This page titled 2.3.6.4: Diagnostic risk assessment approaches and tools is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **6.4: Diagnostic risk assessment approaches and tools** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.6.5: Regulatory Frameworks

6.5. Regulatory Frameworks

Regulatory frameworks

Authors: Charles Bodar and Joop de Knecht

Reviewers: Kees van Gestel

Learning objectives:

You should be able to

- explain how the potential environmental risks of chemicals are legally being controlled in the EU and beyond
- mention the different regulatory bodies involved in the regulation of different categories of chemicals
- explain the purpose of the Classification, Labelling and Packaging (CLP) approach and its difference with the risk assessment of chemicals

Keywords: chemicals, environmental regulations, hazard, risk

Introduction

There is no single, overarching global regulatory framework to manage the risks of all chemicals. Instead, different regulations or directives have been developed for different categories of chemicals. These categories are typically related to the usage of the chemicals. Important categories are industrial chemicals (solvents, plasticizers, etc.), plant protection products, biocides and human and veterinary drugs. Some chemicals may belong to more than one category. Zinc, for example, is used in the building industry, but it also has biocidal applications (antifouling agent) and zinc oxide is used as a veterinary drug. In the European Union, each chemical category is subject to specific regulations or directives providing the legal conditions and requirements to guarantee a safe production and use of chemicals. A key element of all legal frameworks is the requirement that sufficient data on a chemical should be made available. Valid data on production and identity (e.g. chemical structure), use volumes, emissions, environmental fate properties and the (eco)toxicity of a chemical are the essential building blocks for a sound assessment and management of environmental risks. Rules for the minimum data set that should be provided by the actors involved (e.g. producers or importers) are laid down in various regulatory frameworks. With this data, both hazard and risk assessments can be carried out according to specified technical guidelines. The outcome of the assessment is then used for risk management, which is focused on minimizing any risk by taking measures, ranging from requests for additional data to restrictions on the particular use or a full-scale ban of a chemical.

REACH

REACH is a regulation of the European Union, adopted to improve the protection of human health and the environment from the risks that can be posed by chemicals, while enhancing the competitiveness of the EU chemicals industry. REACH stands for Registration, Evaluation, Authorisation and Restriction of Chemicals. The REACH regulation entered into force on 1st June 2007 to streamline and improve the former legislative frameworks on new and existing chemical substances. It replaced approximately forty community regulations and directives by one single regulation.

REACH establishes procedures for collecting and assessing information on the properties, hazards and risks of substances. REACH applies to a very broad spectrum of chemicals: from industrial to household applications, and very much more. It requires that EU manufacturers and importers register their chemical substances if produced or imported in annual amounts of > 1 tonne, unless the substance is exempted from registration under REACH. At quantities of > 10 tonnes, the manufacturers, importers, and down-stream users are responsible to show that their substances do not adversely affect human health or the environment.

The amount of standard information required to show safe use depends on the quantity of the substance that is manufactured or imported. Before testing on vertebrate animals like fish and mammals, the use of alternative methods must be considered. The European Chemical Agency (ECHA) coordinates and facilitates the REACH program. For production volumes above 10 tonnes per year, industry has to prepare a risk assessment, taking into account all risk management measures envisaged, and document this in a chemical safety assessment (CSA). A CSA should include an exposure assessment, hazard or dose-response assessment, and a risk characterization showing risks ratios below 1.0, i.e. safe use (see sections on REACH Human and REACH Eco).

Classification, Labelling and Packaging (CLP)

The EU CLP regulation requires manufacturers, importers or downstream users of substances or mixtures to classify, label and package their hazardous chemicals appropriately before placing them on the market. When relevant information (e.g. ecotoxicity data) on a substance or mixture meets the classification criteria in the CLP regulation, the hazards of a substance or mixture are identified by assigning a certain hazard class and category. An important CLP hazard class is '*Hazardous to the aquatic environment*', which is divided into categories based on the criteria, for example, *Category Acute 1*, representing the most (acute) toxic chemicals (LC50/EC50 ≤ 1 mg/L). CLP also sets detailed criteria for the labelling elements, such as the well-known pictograms (Figure 1).



Figure 1: Pictogram used to indicate hazardousness to the environment. (source: https://www.pictogrammenwinkel.nl/index.php?main_page=product_info&products_id=5988, with permission)

Plant protection products regulation

Plant protection products (PPPs) are pesticides that are mainly used to keep crops healthy and prevent them from being damaged by disease and infestation. They include among others herbicides, fungicides, insecticides, acaricides, plant growth regulators and repellents (see section on Crop Protection Products). PPPs fall under the EU Regulation (EC) No 1107/2009 which determines that PPPs cannot be placed on the market or used without prior authorization. The European Food and Safety Authority (EFSA) coordinates the EU regulation on PPPs.

Biocides regulation

The distinction between biocides and PPP is not always straightforward, but as a general rule of thumb the PPP regulation applies to substances used by farmers for crop protection while the biocides regulation covers all other pesticide applications. Different applications of the same active ingredient, one as a PPP and the other as a biocide, may thus fall under different regulations. Biocides are used to protect humans, animals, materials or articles against harmful organisms like pests or bacteria, by the action of the active substances contained in the biocidal product. Examples of biocides are antifouling agents, preservatives and disinfectants.

According to the EU Biocidal Products Regulation (BPR), all biocidal products require an authorization before they can be placed on the market, and the active substances contained in the biocidal product must be previously approved. The European Chemical Agency (ECHA) coordinates and facilitates the BPR. More or less similar to other legislations, for biocides the environmental risk assessment is mainly performed by comparing compartmental concentrations (PEC) with the concentration below which unacceptable effects on organisms will most likely not occur (PNEC).

Veterinary and human pharmaceuticals regulation

Since 2006, EU law requires an environmental risk assessment (ERA) for all new applications for a marketing authorization of human and veterinary pharmaceuticals. For both products, guidance documents have been developed for conducting an ERA based on two phases. The first phase estimates the exposure of the environment to the drug substance. Based on an action limit the assessment may be terminated. In the second phase, information about the fate and effects in the environment is obtained and assessed. For conducting an ERA a base set, including ecotoxicity data, is required. For veterinary medicines, the ERA is part of a risk-benefit analysis, in which the positive therapeutic effects are weighed against any environment risks, whereas for human medicines the environmental concerns are excluded from the risk-benefit analysis. The European Medicines Agency (EMA) is responsible for the scientific evaluation, supervision and safety monitoring of medicines in the EU.

Harmonization of testing

Testing chemicals is an important aspect of risk assessment, e.g. testing for toxicity, for degradation or for a physicochemical property like the K_{ow} (see Chapter 3). The outcome of a test may vary depending on the conditions, e.g. temperature, test medium or light conditions. For this reason there is an incentive to standardize the test conditions and to harmonize the testing procedures between agencies and countries. This would also avoid duplication of testing, and leading to a more efficient and effective testing system.

The Organization for Economic Co-operation and Development (OECD) assists its member governments in developing and implementing high-quality chemical management policies and instruments. One of the key activities to achieve this goal is the development of harmonized guidelines to test and assess the risks of chemicals leading to a system of mutual acceptance of chemical safety data among OECD countries. The OECD also developed Principles of Good Laboratory Practice (GLP) to ensure that studies are of sufficient quality and rigor and are verifiable. The OECD also facilitates the development of new tools to obtain more safety information and maintain quality while reducing costs, time and animal testing, such as the OECD QSAR toolbox.

6.5. Question 1

What are the major categories of chemicals for which regulatory frameworks are present controlling the environmental risks of these chemicals?

6.5. Question 2

Is a chemical producer in the EU allowed to put a new chemical on the market without a registration or authorization?

6.5. Question 3

Is the CLP regulation based on the hazard, exposure and/or risk of chemicals?

6.5. Question 4

The amount of minimum information that is required under REACH for making a proper hazard and risk assessment is dependent on what? And what do you think is the rationale behind it?

6.5. Question 5

Name three European agencies or authorities that are coordinating important regulatory frameworks in the EU?

6.5.1. REACH human

Authors: Theo Vermeire

Reviewers: Tim Bowmer

Learning objective:

You should be able to:

- outline how human risk assessment of chemicals is performed under REACH;
- explain the regulatory function of human risk assessment in REACH.

Keywords: REACH, chemical safety assessment, human, RCR, DNEL, DMEL

Human risk assessment under REACH

The REACH Regulation aims to ensure a high level of protection of human health and the environment, including the promotion of alternative methods for assessment of hazards of substances, as well as the free circulation of substances on the internal market while enhancing competitiveness and innovation. Risk assessment under REACH aims to realize such a level of protection for humans that the likelihood of adverse effects occurring is low, taking into account the nature of the potentially exposed population (including sensitive groups) and the severity of the effect(s). Industry therefore has to prepare a risk assessment (in REACH terminology: chemical safety assessment, CSA) for all relevant stages in the life cycle of the chemical, taking into account all risk management measures envisaged, and document this in the chemical safety report (CSR). Risk characterization in the context of a CSA is the estimation of the likelihood that adverse effect levels occur due to actual or predicted exposure to a chemical. The human populations considered, or protection goals, are workers, consumers and humans exposed via the environment. In risk

characterization, exposure levels are compared to reference levels to yield "risk characterization ratios" (RCRs) for each protection goal. RCRs are derived for all endpoints (e.g. skin and eye irritation, sensitization, repeated dose toxicity) and time scales. It should be noted that these RCRs have to be derived for all stages in the life-cycle of a compound.

Environmental exposure assessment for humans

Humans can be exposed through the environment directly via inhalation of indoor and ambient air, soil ingestion and dermal contact, and indirectly via food products and drinking water (Figure 1). REACH does not consider direct exposure via soil.

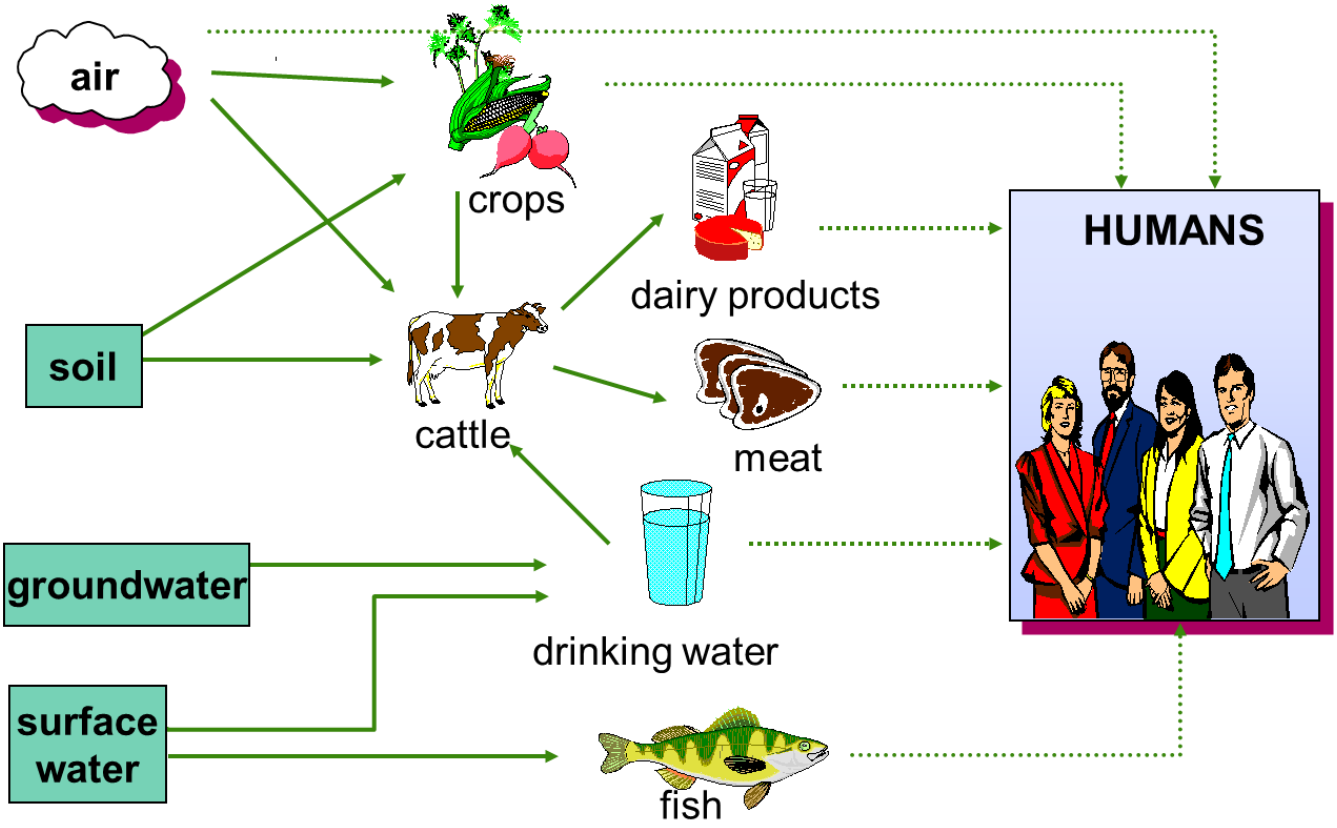


Figure 1. Main exposure routes considered in REACH for environmental exposure of humans.

In the REACH exposure scenario, assessment of human exposure through the environment can be divided into three steps:

1. Determination of the concentrations in intake media (air, soil, food, drinking water);
2. Determination of the total daily intake of these media;
3. Combining concentrations in the media with total daily intake (and, if necessary, using a factor for bioavailability through the route of uptake concerned).

A fourth step may be the consideration of aggregated exposure taking into account exposure to the same substance in consumer products and at the workplace. Moreover, there may be similar substances, acting via the same mechanism of action, that may have to be considered in the exposure assessment, for instance, as a worst case, by applying the concept of dose or concentration addition.

The section on Environmental realistic scenarios (PECs) - Human explains the concept of exposure scenarios and how concentrations in environmental compartments are derived.

Hazard identification and dose-response assessment

The aim of hazard identification is to classify chemicals and to select key data for the dose-response assessment to derive a safe reference level, which in REACH terminology is called the DNEL (Derived No Effect Level) or DMEL (Derived Minimal Effect Level). For human end-points, a distinction is made between substances considered to have a threshold for toxicity and those

without a threshold. For threshold substances, a No-Observed-Adverse Effect Level (NOAEL) or Lowest-Observed-Adverse-Effect Level (LOAEL) is derived, typically from toxicity studies with laboratory animals such as rats and mice. Alternatively a Benchmark Dose (BMD) can be derived by fitting a dose-response model to all observations. These toxicity values are then extrapolated to a DNEL using assessment factors to correct for uncertainty and variability. The most frequently used assessment factors are those for interspecies differences and those for intraspecies variability (see section on Setting safe standards). Additionally, factors can be applied to account for remaining uncertainties such as those due to a poor database.

For substances considered to exert their effect by a non-threshold mode of action, especially mutagenicity and carcinogenicity, it is generally assumed, as a default assumption, that even at very low levels of exposure residual risks cannot be excluded. That said, recent progress has been made on establishing scientific, 'health-based' thresholds for some genotoxic carcinogens. For non-threshold genotoxic carcinogens it is recommended to derive a DMEL, if the available data allow. A DMEL is a cancer risk value considered to be of very low concern, e.g. a 1 in a million tumour risk after lifetime exposure to the chemical and using a conservative linear dose-response model. There is as yet no EU-wide consensus on acceptable levels of cancer risk.

Risk characterization

Safe use of substances is demonstrated when:

• RCRs are below one, both at local and regional level. For threshold substances, the RCR is the ratio of the estimated exposure (concentration or dose) and the DNEL; for non-threshold substances the DMEL is used.

• The likelihood and severity of an event such as an explosion occurring due to the physicochemical properties of the substance as determined in the hazard assessment is negligible.

A risk characterization needs to be carried out for each exposure scenario (see Section on Environmental realistic scenarios (PECs) - Human) and human population. The assessment consists of a comparison of the exposure of each human population known to be

or likely to be exposed with the appropriate DNELs or DMELs and an assessment of the likelihood and severity of an event occurring due to the physicochemical properties of the substance.

Example of a deterministic assessment (Vermeire et al., 2001)

Exposure assessment

Based on an emission estimation for processing of dibutylphthalate (DBP) as a softener in plastics, the concentrations in environmental compartments were estimated. Based on modelling as schematically presented in Figure 1, the total human dose was determined to be 93 ug.kg $_{bw}^{-1}$.

PEC-air	2.4	µg.m^{-3}
PEC-surface water	2.8	µg.l^{-1}
PEC- grassland soil	0.15	mg.kg^{-1}
PEC-porewater agric. soil	3.2	µg.l^{-1}
PEC-porewater grassl. soil	1.4	µg.l^{-1}
PEC-groundwater	3.2	µg.l^{-1}
Total Human Dose	93	µg.kg$_{bw}^{-1}$.d^{-1}

Effects assessment

The total dose should be compared to a DNEL for humans. DBP is not considered a genotoxic carcinogen but is toxic to reproduction and therefore the risk assessment is based on endpoints assumed to have a threshold for toxicity. The lowest NOAEL of DBP was observed in a two-generation reproduction test in rats and at the lowest dose-level in the diet (52 mg.kg$_{bw}^{-1}$.d^{-1} for males and 80 mg.kg$_{bw}^{-1}$.d^{-1} for females) a reduced number of live pups per litter and decreased pup weights were seen in the absence of maternal toxicity. The lowest dose level of 52 mg.kg$_{bw}^{-1}$.d^{-1} was chosen as the NOAEL. The DNEL was derived by the application of an overall assessment factor of 1000, accounting for interspecies differences, human variability and uncertainties due to a non-chronic exposure period.

Risk characterisation

The deterministic estimate of the RCR would be based on the deterministic exposure estimate of 0.093 mg.kg$_{bw}^{-1}$.d^{-1} and the deterministic DNEL of 0.052 mg.kg$_{bw}^{-1}$.d^{-1}. The deterministic RCR would then be 1.8, based on the NOAEL. Since this is higher than one, this assessment indicates a concern, requiring a refinement of the assessment or risk management measures.

Additional reading

Van Leeuwen C.J., Vermeire T.G. (Eds.) (2007) Risk assessment of chemicals: an introduction. Springer, Dordrecht, The Netherlands, ISBN 978-1-4020-6102-8 (e-book), https://doi.org/10.1007/978-1-4020-6102-8.

Vermeire, T., Jager, T., Janssen, G., Bos, P., Pieters, M. (2001) A probabilistic human health risk assessment for environmental exposure to dibutylphthalate. Journal of Human and Ecological Risk Assessment 7, 1663-1679.

6.5.1. Question 1

Uncertainty happens! It is inherent to risk assessment. Where, in your view, are the greatest sources of uncertainty in the process of risk assessment?

6.5.1. Question 2

Are there risks identified in the example for humans indirectly exposed via the environment? To what extent are these potential or realistic risks by asking yourself the following questions:

• Do you have relevant toxicological and exposure data ?

• Are these fixed values or not?

• How relevant or adverse are the toxicological effects observed?

• Were appropriate assessment factors used?

What would you recommend as a strategy to reduce the identified risks sufficiently?

6.5.2. REACH environment

Author: Joop de Knecht

Reviewers: Watze de Wolf

Keywords: REACH, European chemicals regulation

Introduction

REACH establishes procedures for collecting and assessing information on the properties, hazards and risks of substances. At quantities of > 10 tonnes, the manufacturers, importers, and down-stream users must show that their substances do not adversely affect human health or the environment for the uses and operational conditions registered. The amount of standard information required to show safe use depends on the quantity of the substance that is manufactured or imported. This section explains how risks to the environment are assessed in REACH.

Data requirements

As a minimum requirement, all substances manufactured or imported in quantities of 1 tonne or more need to be tested in acute toxicity tests on *Daphnia* and algae, while also information should be provided on biodegradability (Table 1). Physical-chemical properties relevant for environmental fate assessment that have to be provided at this tonnage level are water solubility, vapour pressure and octanol-water partition coefficient. At 10 tonnes or more, this should be supplemented with an acute toxicity test on fish and an activated sludge respiration inhibition test. At this tonnage level, also an adsorption/desorption screening and a hydrolysis test should be performed. If the chemical safety assessment, performed at 100 tonnes or more in case a substance is classified based on hazard information, indicates the need to investigate further the effects on aquatic organisms, the chronic toxicity on these aquatic species should be determined. If the substance has a high potential for bioaccumulation (for instance a log K_{ow} > 3), also the bioaccumulation in aquatic species should be determined. The registrant should also determine the acute toxicity to terrestrial species or, in the absence of these data, consider the equilibrium partitioning method (EPM) to assess the hazard to soil organisms. To further investigate the fate of the substance in surface water, sediment and soil, simulation tests on its degradation should be conducted and when needed further information on the adsorption/desorption should be provided. At 1000 tonnes or more, chronic tests on terrestrial and sediment-living species should be conducted if further refinement of the safety assessment is needed. Before testing vertebrate animals like fish and mammals, the use of alternative methods and all other options must be considered to comply with the regulations regarding (the reduction of) animal testing.

Table 1 Required ecotoxicological and environmental fate information as defined in REACH

1-10 t/y	• Acute Aquatic toxicity (invertebrates, algae) • Ready biodegradability
10-100 t/y	• Acute Aquatic toxicity (fish) • Activated sludge respiration, inhibition • Hydrolysis as a function of pH • Adsorption/ desorption screening test
100-1000 t/y	• Chronic Aquatic toxicity (invertebrates, fish) • Bioaccumulation • Surface water, soil and sediment simulation (degradation) test • Acute terrestrial toxicity • Further information on adsorption/desorption
≥ 1000 t/y	• Further fate and behaviour in the environment of the substance and/or degradation products • Chronic terrestrial toxicity • Sediment toxicity • Avian toxicity

Safety assessment

For substances that are classified based on hazard information the registrant should assess the environmental safety of a substance, by comparing the predicted environmental concentration (PEC) with the Predicted No Effect Concentration (PNEC), resulting in a Risk Characterisation Ratio (RCR=PEC/PNEC). The use of the substance is considered to be safe when the RCR <1.

Chapter 16 of the ECHA guidance offers methods to estimate the PEC based on the tonnage, use and operational conditions, standardised through a set of use descriptors, particularly the Environmental Release Categories (ERCs). These ERCs are linked to conservative default release factors to be used as a starting point for a first tier environmental exposure assessment. When substances are emitted via waste water, the physical-chemical and fate properties of the chemical substance are then used to predict its behaviour in the Wastewater Treatment Plant (WWTP). Subsequently, the release of treated wastewater is used to estimate the concentration in fresh and marine surface water. The concentration in sediment is estimated from the PEC in water and experimental or estimated sediment-water partitioning coefficient (Kp_{sed}). Soil concentrations are estimated from deposition from air and the application of sludge from an WWTP. The guidance offers default values for all relevant parameters, thus a generic local PEC can be calculated and considered applicable to all local emissions in Europe, although the default values can be adapted to specific conditions if justified. The local risk for wide-dispersive uses (e.g. from consumers or small, non- industrial companies) is estimated for a default WWTP serving 10,000 inhabitants. In addition, a regional assessment is conducted for a standard area, a region represented by a typical densely populated EU area located in Western Europe (i.e. about 20 million inhabitants, distributed in a 200 x 200 km^2 area). For calculating the regional PECs, a multi-media fate-modelling approach is used (e.g. the SimpleBox model; see Section on Multicompartment fate modelling). All releases to each environmental compartment for each use, assumed to constitute a constant and continuous flux, are summed and averaged over the year and steady-state concentrations in the environmental compartments are calculated. The regional concentrations are used as background concentrations in the calculation of the local concentrations.

The PNEC is calculated using the lowest toxicity value and an assessment factor (AF) related to the amount of information (see section Setting safe standards or chapter 10 of the REACH guidance. If only the minimum set of aquatic acute toxicity data is available, i.e. LC50s or EC50s for algae, daphnia and fish, a default value of 1000 is used. When one, two or three or more long-term tests are available, a default AF of 100, 50 and 10 is applied to No Observed Effect Concentrations (NOECs), respectively. The idea behind lowering the AF when more data become available is that the amount of uncertainty around the PNEC is being reduced.

In the absence of ecotoxicological data for soil and/or sediment-dwelling organisms, the $PNEC_{soil}$ and/or $PNEC_{sed}$ may be provisionally calculated using the EPM. This method uses the $PNEC_{water}$ for aquatic organisms and the suspended matter/water

partitioning coefficient as inputs. For substances with a log K_{ow} >5 (or with a corresponding log Kp value), the PEC/PNEC ratio resulting from the EPM is increased by a factor of 10 to take into account possible uptake through the ingestion of sediment. If the PEC/PNEC is greater than 1 a sediment test must be conducted. If one, two or three long-term No Observed Effect Concentrations (NOECs) from sediment invertebrate species representing different living and feeding conditions are available, the PNEC can be derived using default AFs of 100, 50 and 10, respectively.

For data rich chemicals, the PNEC can be derived using Species Sensitivity Distributions (SSD) or other higher-tier approaches.

6.5.2. Question 1

Who is responsible within the EU that industrial chemicals do not pose a risk for the environment?

6.5.2. Question 2

In which circumstances will an environmental risk be identified?

6.5.2. Question 3

Describe how the ecotoxicogical safety levels (PNECs) for the aquatic environment are derived depending on the ecotoxicological information available?

6.5.3. Pesticides (EFSA)

under review

6.5.4. Environmental Risk Assessment of Pharmaceuticals in Europe

Author: Gerd Maack

Reviewers: Ad Ragas, Julia Fabrega, Rhys Whomsley

Learning objectives:

You should be able to

- Explain the philosophy and objective of the environmental risk assessment of pharmaceuticals.
- mention the key aspects of the tiered approach of the assessment
- identify the exposure routes for human and veterinary medicinal products and should know the respective consequences in the assessment

Keywords: Human pharmaceuticals, veterinary pharmaceuticals, environmental impact, tiered approach

Introduction

Pharmaceuticals are a crucial element of modern medicine and confer significant benefits to society. About 4,000 active pharmaceutical ingredients are being administered worldwide in prescription medicines, over-the-counter medicines, and veterinary medicines. They are designed to be efficacious and stable, as they need to pass different barriers i.e. skin, the gastrointestinal system (GIT), or even the blood-brain barrier before reaching the target cells. Each target system has a different pH and different lipophilicity and the GIT is in addition colonised with specific bacteria, specialized to digest, dissolve and disintegrate organic molecules. As a consequence of this stability, most of the pharmaceutical ingredients are stable in the environment as well and could cause effects on non-target organisms.

The active ingredients comprise a variety of synthetic chemicals produced by pharmaceutical companies in both the industrialized and the developing world at a rate of 100,000 tons per year.

While pharmaceuticals are stringently regulated in terms of efficacy and safety for patients, as well as for target animal safety, user and consumer safety, the potential effects on non-target organisms and environmental effects are regulated comparably weakly.

The authorisation procedure requires an environmental risk assessment (ERA) to be submitted by the applicants for each new human and veterinary medicinal product. The assessment encompasses the fate and behaviour of the active ingredient in the environment and its ecotoxicity based on a catalogue of standardised test guidelines.

In the case of veterinary pharmaceuticals, constraints to reduce risk and thus ensure safe usage can be stipulated in most cases. In the case of human pharmaceuticals, it is far more difficult to ensure risk reduction through restriction of the drug's use due to practical and ethical reasons. Because of their unique benefits, a restriction is not reasonable. This is reflected in the legal framework, as a potential effect on the environment is not included in the final benefit risk assessment for a marketing authorisation.

Exposure pathways

Human pharmaceuticals

Human pharmaceuticals enter the environment mainly via surface waters through sewage systems and sewage treatment plants. The main exposure pathways are excretion and non-appropriate disposal. Typically only a fraction of the medicinal product taken is metabolised by the patients, meaning that the main share of the active ingredient is excreted unchanged into the wastewater system. Furthermore, sometimes the metabolites themselves are pharmacologically active. Yet, no wastewater treatment plant is able to degrade all active ingredients. So medicinal products are commonly found in surface water, to some extent in ground water, and sometimes even in drinking water. However, the concentrations in drinking water are orders of magnitude lower than therapeutic concentrations. An additional exposure pathway for human pharmaceuticals is the spreading of sewage sludge on soil, if the sludge is used as fertilizer on farmland. See for more details the Link "*The Drugs We Wash Away: Pharmaceuticals, Drinking Water and the Environment*".

Veterinary pharmaceuticals

Veterinary pharmaceuticals on the other hand enter the environment mainly via soil, either indirectly, if the slurry and manure from mass livestock production is spread onto agricultural land as fertiliser, or directly from pasture animals. Moreover, pasture animals might additionally excrete directly into surface water. Pharmaceuticals can also enter the environment via the detour of manure used in biogas plants.

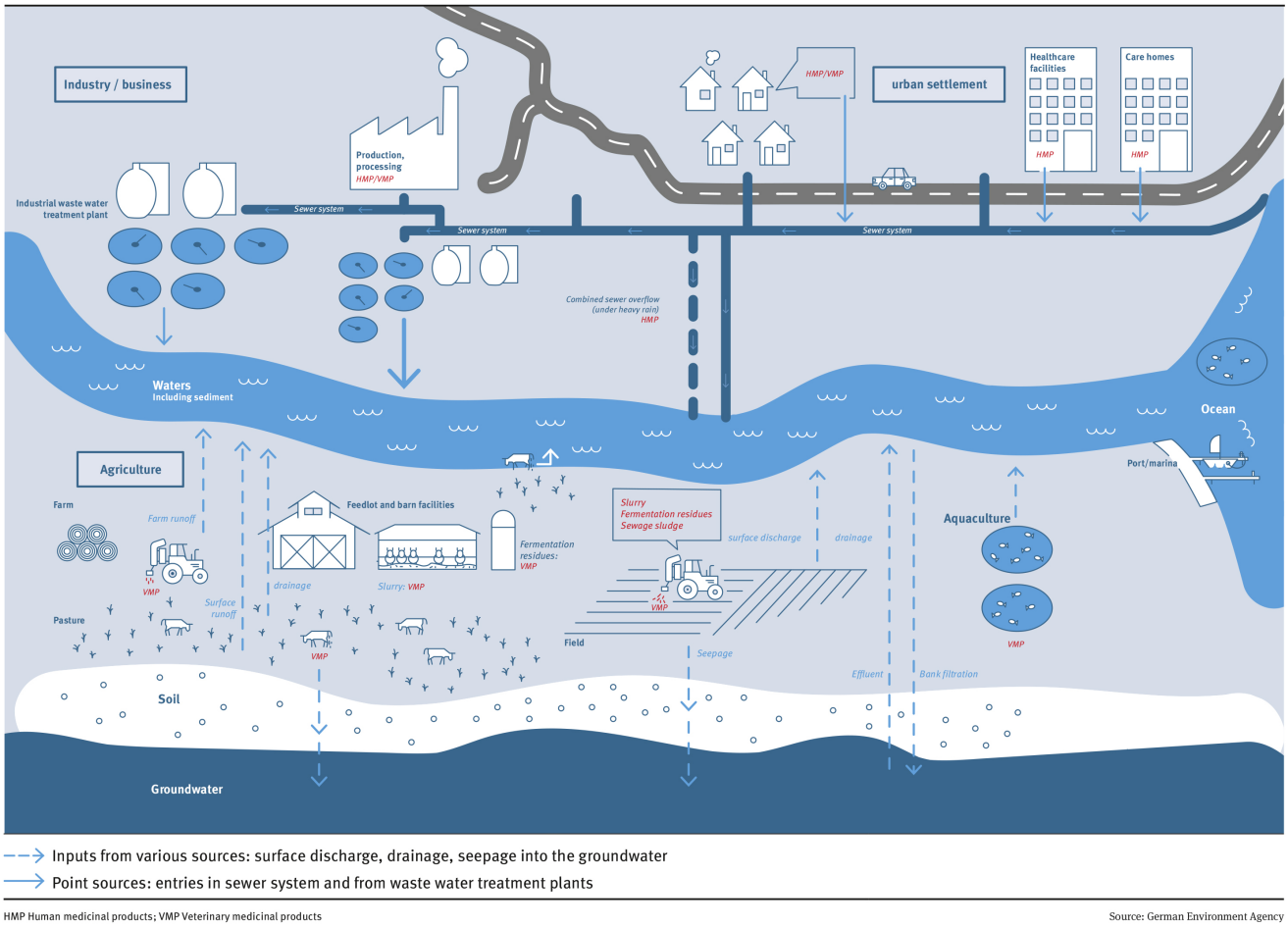


Figure 1: Entry path paths of human and veterinary medicinal products. See text for more details (reproduced with permission from the German Environmental Agency).

Assessment schemes

Despite the differences mentioned above, the general scheme of the environmental risk assessment of human and veterinary pharmaceuticals is similar. Both assessments start with an exposure assessment. Only if specific trigger values are reached an in-depth assessment of fate, behaviour and effects of the active ingredient is necessary.

Environmental risk assessment of human pharmaceuticals

In Europe, an ERA for human pharmaceuticals has to be conducted according to the *Guideline on Environmental Risk Assessment of Medicinal Products for Human Use* (EMA 2006). This ERA consists of two phases. Phase I is a pre-screening for estimating the exposure in surface water, and if this Predicted Environmental Exposure Concentration (PEC) does not reach the action limit of 0.01 µg/L, in most cases, the ERA can stop. In case this action limit is reached or exceeded, a base set of aquatic toxicology, and fate and behaviour data need to be supplied in phase II Tier A. A risk assessment, comparing the PEC with the Predicted No Effect Concentration (PNEC), needs to be conducted. If in this step a risk is still identified for a specific compartment, a substance and compartment-specific refinement and risk assessment in Phase II Tier B needs to be conducted (Figure 2).

Phase I: Estimation of Exposure

In Phase I, the PEC calculation is restricted to the aquatic compartment. The estimation should be based on the drug substance only, irrespective of its route of administration, pharmaceutical form, metabolism and excretion. The initial calculation of the PEC in surface water assumes:

- The predicted amount used per capita per year is evenly distributed over the year and throughout the geographic area ($Dose_{ai}$);
- A fraction of the overall market penetration (F_{pen}), in other words 'how many people will take the medicinal product? Normally a default value of 1% is used;
- The sewage system is the drug's main route of entry into surface water.

The following formula is used to estimate the PEC in surface water:

$$PEC_{surfacewater} = \frac{(DOSE_{ai} * F_{pen})}{(WASTE_{inhab} * DILUTION)}$$

$PEC_{surfacewater}$ = mg/l

$DOSE_{ai}$ = Maximum daily dose consumed per capita [$mg \cdot inh^{-1} \cdot d^{-1}$]

F_{pen} = Fraction of market penetration (= 1% by default)

$WASTE_{inhab}$ = Amount of wastewater per inhabitant per day (= 200 l by default)

DILUTION = Dilution Factor (= 10 by default)

Three factors of this formula, i.e. F_{pen}, $Waste_{inhab}$ and the Dilution Factor, are default values, meaning that the $PEC_{surfacewater}$ in Phase I entirely depends on the dose of the active ingredient. The Fpen can be refined by providing reasonably justified market penetration data, e.g. based on published epidemiological data.

If the $PEC_{surfacewater}$ value is equal to or above 0.01 µg/l (mean dose ≥ 2 mg cap^{-1} d^{-1}), a Phase II environmental fate and effect analysis should be performed. Otherwise, the ERA can stop. However, in some cases, the action limit may not be applicable. For instance, medicinal substances with a log K_{ow} > 4.5 are potential PBT candidates and should be screened for persistence (P), bioaccumulation potential (B), and toxicity (T) independently of the PEC value. Furthermore, some substances may affect vertebrates or lower animals at concentrations lower than 0.01 µg/L. These substances should always enter Phase II and a tailored risk assessment strategy should be followed which addresses the specific mechanism of action of the substance. This is often true for e.g. hormone active substances (see section on Endocrine disruption). The required tests in a Phase II assessment (see below) need to cover the most sensitive life stage, and the most sensitive endpoint needs to be assessed. This means for instance that for substances affecting reproduction, the organism needs to be exposed to the substance during gonad development and the reproductive output needs to be assessed.

Phase II: Environmental Fate and Effects Analysis

A Phase II assessment is conducted by evaluating the PEC/PNEC ratio based on a base set of data and the predicted environmental concentration from Tier A. If a potential environmental impact is indicated, further testing might be needed to refine PEC and PNEC values in Tier B.

Under certain circumstances, effects on sediment-dwelling organisms and terrestrial environmental fate and effects analysis are also required. Experimental studies should follow standard test protocols, e.g. OECD guidelines. It is not acceptable to use QSAR estimation, modelling and extrapolation from e.g. a substance with a similar mode of action and molecular structure (read across). This is in clear contrast to other regulations like e.g. REACH.

Human pharmaceuticals are used all year round without any major fluctuations and peaks. The only exemption are substances used against cold and influenza. These substances have a clear peak in the consumption in autumn and winter times. In developed countries in Europe and North America, antibiotics display a similar peak as they are prescribed to support the substances used against viral infections. The guideline reflects this exposure scenario and asks explicitly for long-term effect tests for all three trophic levels: algae, aquatic invertebrates and vertebrates (i.e., fish).

In order to assess the physio chemical fate, amongst other tests the sorption behaviour and fate in a water/sediment system should be determined.

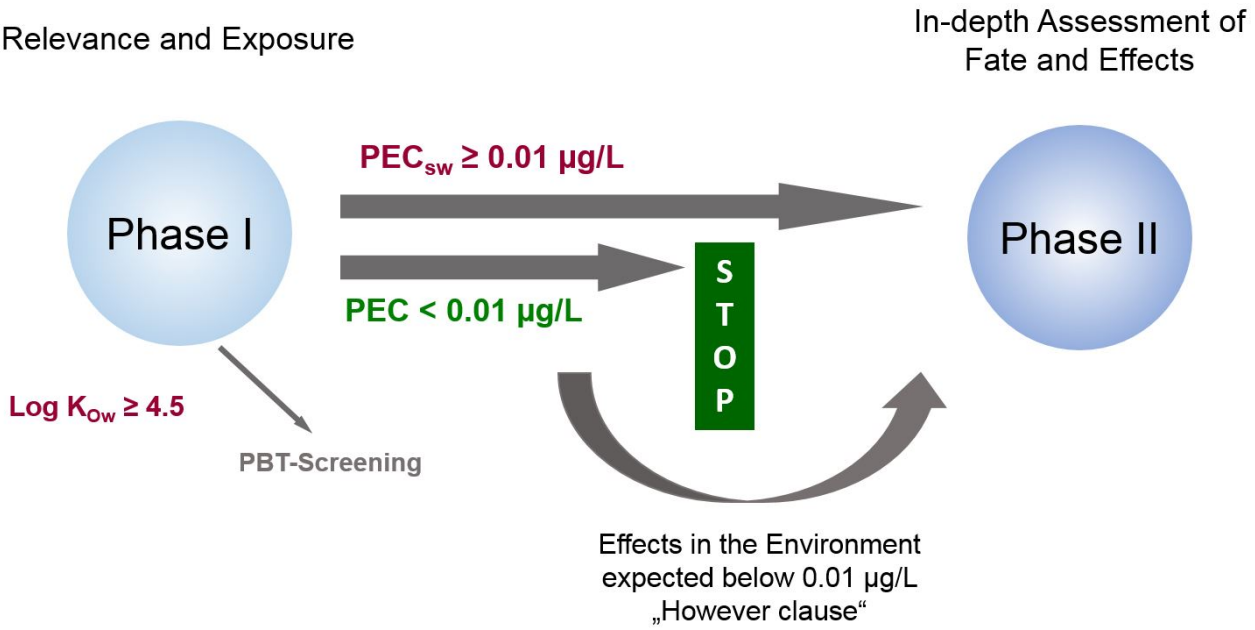


Figure 2: Scheme of conducting an ERA for Human Medicinal Products according to the EMA guideline

If, after refinement, the possibility of environmental risks cannot be excluded, precautionary and safety measures may consist of:

- An indication of potential risks presented by the medicinal product for the environment.
- Product labelling, Summary Product Characteristics (SPC), Package Leaflet (PL) for patient use, product storage and disposal.

Labelling should generally aim at minimising the quantity discharged into the environment by appropriate mitigation measures

Environmental risk assessment of veterinary pharmaceuticals

In the EU, the Environmental Risk Assessment (ERA) is conducted for all veterinary medicinal products. The structure of an ERA for Veterinary Medicinal Products (VMPs) is quite similar to the ERA for Human Medicinal Products. It is also tier based and starts with an exposure assessment in Phase I. Here, the potential for environmental exposure is assessed based on the intended use of the product. It is assumed that products with limited environmental exposure will have negligible environmental effects and thus can stop in Phase I. Some VMPs that might otherwise stop in Phase I as a result of their low environmental exposure, may require additional hazard information to address particular concerns associated with their intrinsic properties and use. This approach is comparable to the assessment of Human Pharmaceutical Products, see above.

Phase I: Estimation of Environmental Exposure

For the exposure assessment, a decision tree was developed (Figure 3). The decision tree consists of a number of questions, and the answers of the individual questions will conclude in the extent of the environmental exposure of the product. The goal is to determine if environmental exposure is sufficiently significant to consider if data on hazard properties are needed for characterizing a risk. Products with a low environmental exposure are considered not to pose a risk to the environment and hence these products do not need further assessment. However, if the outcome of Phase I assessment is that the use of the product leads to significant environmental exposure, then additional environmental fate and effect data are required. Examples for products with a low environmental exposure are, among others are products for companion animals only and products that result in a Predicted Environmental Concentration in soil (PECsoil) of less than 100 µg/kg, based on a worst-case estimation.

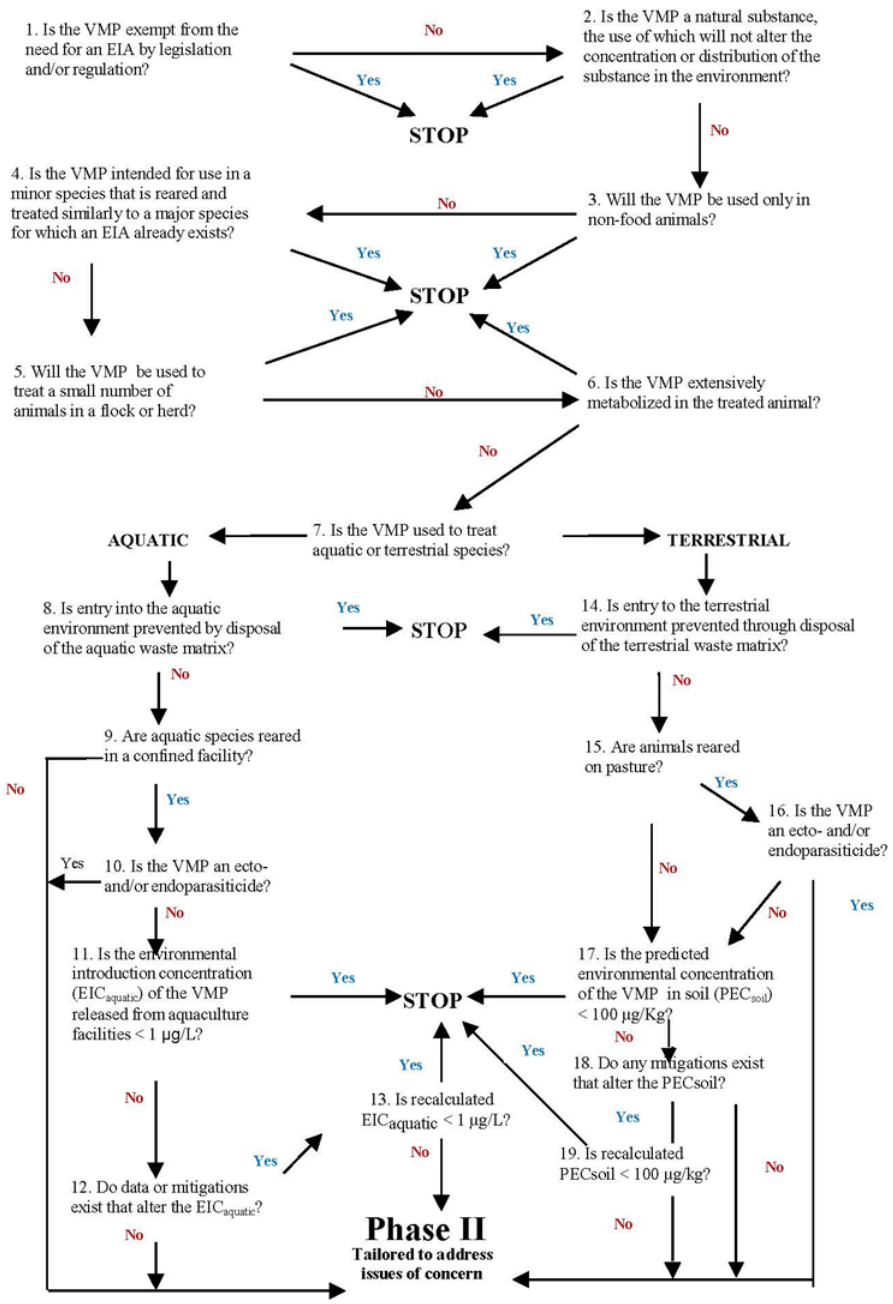


Figure 3: Phase I Decision Tree for Veterinary Medicinal Products (VMPs); (VICH 2000)

Phase II: Environmental Fate and Effects Analysis

A Phase II assessment is necessary if either the trigger of 100 µg/kg in the terrestrial branch or the trigger of 1 µg/L in the aquatic branch is reached. It is also necessary, if the substance is a parasiticide for food producing animals. A Phase II is also required for substances that would in principle stop in Phase I, but there are indications that an environmental risk at very low concentrations is likely due to their hazardous profile (e.g., endocrine active medicinal products). This is comparable to the assessment for Human Pharmaceutical Products.

For Veterinary Pharmaceutical Products also the Phase II assessment is sub-divided into several Tiers, see Figure 4. For Tier A, a base set of studies assessing the physical-chemical properties, the environmental fate, and effects of the active ingredient is necessary. For Tier A, acute effect tests are suggested, assuming a more peak like exposure scenario due to e.g. applying manure and dung on fields and meadows, in contrast to the permanent exposure of human pharmaceuticals. If for a specific trophic level,

e.g. dung fauna or algae, a risk is identified (PEC/PNEC ≥1) (see Introduction to Chapter 6), long-term tests for this level have to be conducted in Tier B. For the trophic levels, without an identified risk, the assessment can stop. If the risk still applies with these long-term studies, a further refinement with field studies in Tier C can be conducted. Here a co-operation with a competent authority is strongly recommended, as these tests are tailored, reflected by the individual design of these field studies. In addition, and independent of this, risk mitigation measures can be imposed to reduce the exposure concentration (PEC). These can be, beside others, that animals must remain stabled for a certain amount of time after the treatment, to ensure that the concentration of active ingredient in excreta is low enough to avoid adverse effects on dung fauna and their predators. Alternatively, the treated animals are denied access to water as the active ingredient has harmful effects on aquatic organisms.

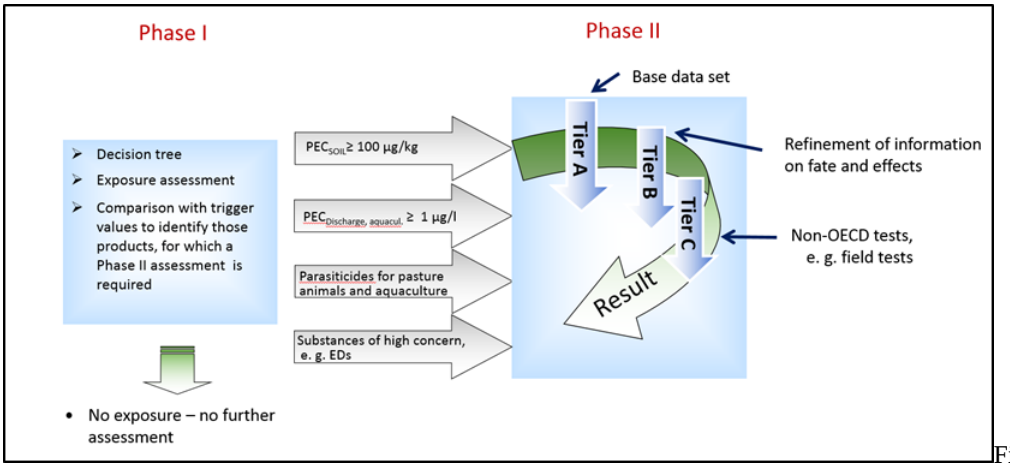


Figure 4: Scheme for conducting an ERA for Veterinary Medicinal Products (VMPs) according to the EMA guidelines (VICH 2000; VICH 2004).

Conclusion

The Environmental Risk Assessment of Human and Veterinary Medicinal Products is a straightforward, tiered-based process with the possibility to exit at several steps in the assessment procedure. Depending on the dose, the physico-chemical properties, and the anticipated use, this can be quite early in the procedure. On the other hand, for very potent substances with specific modes of action the guidelines are flexible enough to allow specific assessments covering these modes of action.

The ERA guideline for human medicinal products entered into application 2006 and many data gaps exist for products approved prior to 2006. Although there is a legal requirement for an ERA dossier for all marketing authorisation applications, new applications for pharmaceuticals on the market before 2006 are only required to submit ERA data under certain circumstances (e.g. significant increase in usage). Even for some of the blockbusters, like Ibuprofen, Diclofenac, and Metformin, full information on fate, behaviour and effects on non-target organisms is currently lacking.

Furthermore, systematic post-authorisation monitoring and evaluation of potential unintended ecotoxicological effects does not exist. The market authorisation for pharmaceuticals does not expire, in contrast to e.g. an authorisation of pesticides, which needs to be renewed every 10 years.

For Veterinary Medicinal Products, an in-depth ERA is necessary for food producing animals only. An ERA for non-food animals can stop with question 3 in Phase I (Figure 3) as it is considered that the use of products for companion animals leads to negligible environmental concentrations, which might not be necessarily the case. Here, the guideline does not reflect the state of the art of scientific and regulatory knowledge. For example, the market authorisation, as a pesticide or biocide, has been withdrawn or strongly restricted for some potent insecticides like imidacloprid and fipronil which both are authorised for use in companion animals.

Further Reading

Pharmaceuticals in the Environment:https://www.umweltbundesamt.de/en/publikationen/pharmaceuticals-in-the-environment-the-global

Recommendations for reducing micro-pollutants in waters: https://www.umweltbundesamt.de/publikationen/recommendations-for-reducing-micropollutants-in

6.5.4. Question 1

Why are pharmaceuticals a problem for non-target organisms and for the environment?

6.5.4. Question 2

How do Human pharmaceuticals enter the environment?

6.5.4. Question 3

How do Veterinary pharmaceuticals enter the environment?

6.5.4. Question 4

What is the general scheme of the environmental risk assessment of human and veterinary pharmaceuticals?

6.5.4. Question 5

Why are long-term tests needed for the assessment of human pharmaceuticals, in contrast to the assessment of veterinary pharmaceuticals?

6.5.5. Water Framework Directive

under review

6.5.6. Policy on soil and groundwater regulation

Author: Frank Swartjes

Reviewers: Kees van Gestel, Ad Ragas, Dietmar Müller-Grabherr

Learning objectives:

You should be able to

- explain how different countries regulate soil contamination issues
- list some differences between different policy systems on soil and groundwater regulations
- describe how risk assessment procedures are implemented in policy

Keywords: Policy on soil contamination, Water Framework Directive, screening values comparison, Thematic Soil Strategy, Common Forum

History

As a bomb hit, soil contamination came onto the political agenda in the United States and in Europe through a number of disasters in the late 1970s and early 1980s. Starting point was the 1978 Love Canal disaster in upper New York State, USA, in which a school and a number of residences had been built on a former landfill for chemical waste disposal with thousands of tonnes of dangerous chemical wastes, and became a national media event. In Europe in 1979, the residential site of Lekkerkerk in the Netherlands became an infamous national event. Again, a residential area had been built on a former waste dump, which included chemical waste from the painting industry, and with channels and ditches that had been filled in with chemical waste-containing materials.

Since these events, soil contamination-related policies emerged one after the other in different countries in the world. Crucial elements of these policies were a benchmark date for a ban on bringing pollutants in or on the soil ('prevention'), including a strict policy, e.g. duty of care, for contaminations that are caused after the benchmark date, financial liability for polluting activities, tools for assessing the quality of soil and groundwater, and management solutions (remediation technologies and facilities for disposal).

Evolution in soil policies

Objectives in soil policies often show evolution over time and changes go along with developing new concepts and approaches for implementing policies. In general, soil policies often develop from a maximum risk control until a functional approach. The corresponding tools for implementation usually develop from a set of screening values towards a systemic use of frameworks,

enabling sound environmental protection while improving the cost-benefit-balance. Consequently, soil policy implementation usually goes through different stages. In general terms, four different stages can be distinguished, i.e., related to maximum risk control, to the use of screening values, to the use of frameworks and based on a functional approach. *Maximum risk control* follows the precaution principle and is a stringent way of assessing and managing contamination by trying to avoid any risk. Procedures based on *screening values* allow for a distinction in polluted and non-polluted sites for which the former, the polluted sites, require some kind of intervention. The scientific underpinning of the earliest generations of screening values was limited and expert judgement played an important role. Later, more sophisticated screening values emerged, based on risk assessment. This resulted in screening values for individual contaminants within the contaminant groups metals and metalloids, other inorganic contaminants (e.g., cyanides), polycyclic aromatic hydrocarbons (PAHs), monocyclic aromatic hydrocarbons (including BETX (benzene, toluene, xylene)), persistent organic pollutants (including PCBs and dioxins), volatile organic contaminants (including trichloroethylene, tetrachloroethylene, 1,1,1-trichloroethane, and vinyl chloride), petroleum hydrocarbons and, in a few countries only, asbestos. For some contaminants such as PAHs, sum-screening values for groups were derived in several countries, based on toxicity equivalents. In a procedure based on *frameworks*, often the same screening values generally act as a trigger for further, more detailed site-specific investigations in one or two additional assessment steps. In the *functional approach*, soil and groundwater must be suited for the land use it relates to (e.g., agricultural or residential land) and the functions (e.g., drinking water abstraction, irrigation water) it performs. Some countries skip the maximum risk control and sometimes also the screening values stages and adopt a framework and/or a functional approach.

European collaboration and legislation

In Europe, collaboration was strengthened by concerted actions such as CARACAS (concerted action on risk assessment for contaminated sites in the European Union; 1996 - 1998) and CLARINET (Contaminated Land Rehabilitation Network for Environmental Technologies; 1998 - 2001). These concerted actions were followed up by fruitful international networks that are still are active today. These are the Common Forum, which is a network of contaminated land policy makers, regulators and technical advisors from Environment Authorities in European Union member states and European Free Trade Association countries, and NICOLE (Network for Industrially Co-ordinated Sustainable Land Management in Europe), which is a leading forum on industrially co-ordinated sustainable land management in Europe. NICOLE is promoting co-operation between industry, academia and service providers on the development and application of sustainable technologies.

In 2000, the EU Water Framework Directive (WFD; Directive 2000/60/EC) was adopted by the European Commission, followed by the Groundwater Directive (Directive 2006/118/EC) in 2006 (European parliament and the council of the European Union, 2019b). The environmental objectives are defined by the WFD. Moreover, 'good chemical status' and the 'no deterioration clause' account for groundwater bodies. 'Prevent and limit' as an objective aims to control direct or indirect contaminant inputs to groundwater, and distinguishes for 'preventing hazardous substances' to enter groundwater as well as 'limiting other non-hazardous substances'. Moreover, the European Commission adopted a Soil Thematic Strategy, with soil contamination being one out of the seven identified threats. A proposal for a Soil Framework Directive, launched in 2006, with the objective to protect soils across the EU, was formally withdrawn in 2014 because of a lack of support from some countries.

Policies in the world

Today, most countries in Europe and North America, Australia and New Zealand, and several countries in Asia and Middle and South America, have regulations on soil and groundwater contamination. The policies, however, differ substantially in stage, extent and format. Some policies only cover prevention, e.g., blocking or controlling the inputs of chemicals onto the soil surface and in groundwater bodies. Other policies cover prevention, risk based quality assessment and risk management procedures and include elaborated technical tools, which enable a decent and uniform approach. In particular the larger countries such as the USA, Germany and Spain, policies differ between states or provinces within the country. And even in countries with a policy on the federal level, the responsibilities for different steps in the soil contamination chain are very different for the different layers of authorities (at the national, regional and municipal level).

In Figure 1 the European countries are shown that have a procedure based on frameworks (as described above), including risk-based screening values. It is difficult, if not impossible, to summarise all policies on soil and groundwater protection worldwide. Alternatively, some general aspects of these policies are given here. A fair first basic element in nearly all soil and groundwater policies, relating to prevention of contamination, is the declaration of a formal point in time after which polluting soil and

groundwater is considered an illegal act. For soil and groundwater quality assessment and management, most policies follow the risk-based land management approach as the ultimate form of the functional approach described above. Central in this approach are the risks for specific targets that need to be protected up to a specified level. Different protection targets are considered. Not surprisingly, 'human health' is the primary protection target that is adopted in nearly all countries with soil and groundwater regulations. Moreover, the ecosystem is an important protection target for soil, while for groundwater the ecosystem as a protection target is under discussion. Another interesting general characteristic of mature soil and groundwater policies is the function-specific approach. The basic principle of this approach is that land must be suited for its purpose. As a consequence, the appraisal of a contaminated site in a residential area, for instance, follows a much more stringent concept than that of an industrial site.

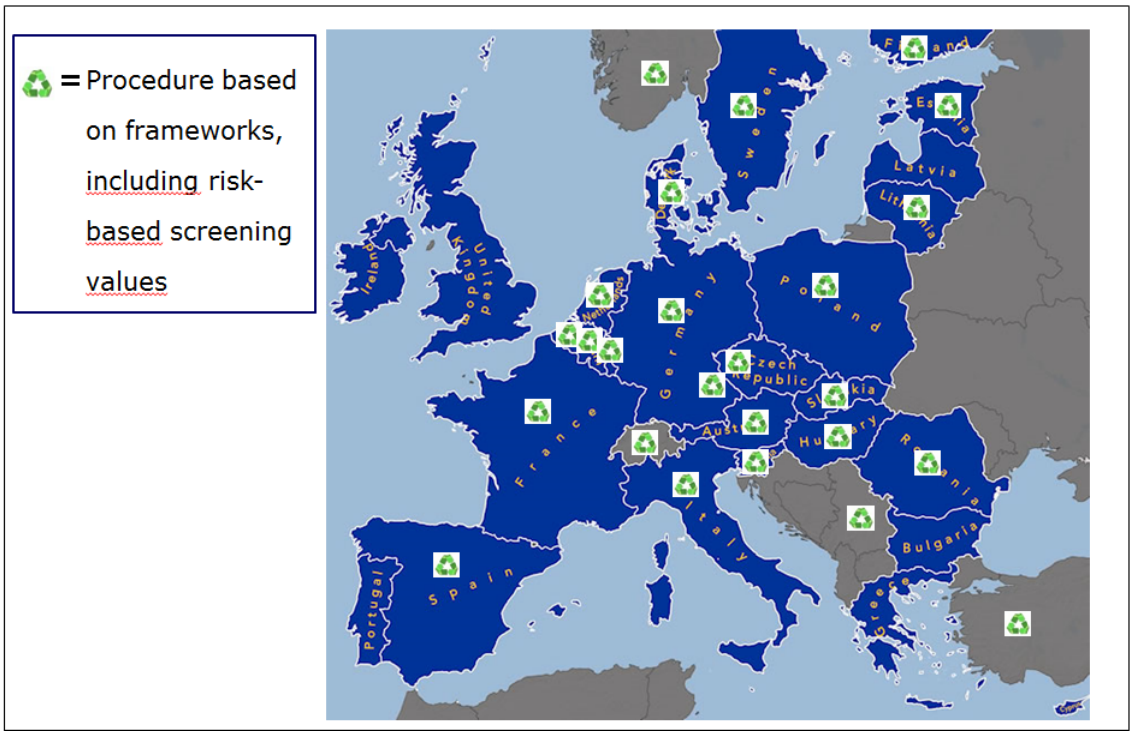


Figure prepared by Frank Swartjes.

Risk assessment tools

Risk assessment tools often form the technical backbone of policies. Since the late 1980s risk assessment procedures for soil and groundwater quality appraisal were developed. In the late 1980s the exposure model CalTOX was developed by the Californian Department of Toxic Substances Control in the USA, a few years later the CSOIL model in the Netherlands (Van den Berg, 1991/ 1994/ 1995). In Figure 2, the flow chart of the Dutch CSOIL exposure model is given as an example. Three elements are recognized in CSOIL, like in most exposure models: (1) contaminant distribution over the soil compartments; (2) contaminant transfer from (the different compartments of) the soil into contact media; and (3) direct and indirect exposure to humans. The major exposure pathways are exposure through soil ingestion, crop consumption and inhalation of indoor vapours (Elert et al., 2011). Today, several exposure models exist (see Figure 3 for some 'national' European exposure models). However, these exposure models may give quite different exposure estimates for the same exposure scenario (Swartjes, 2007).

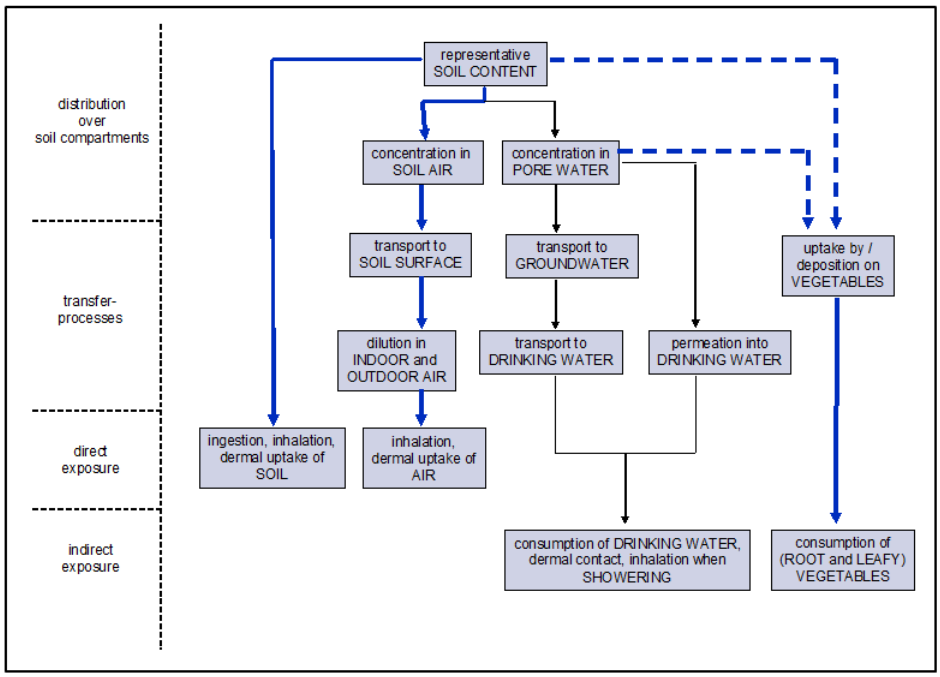
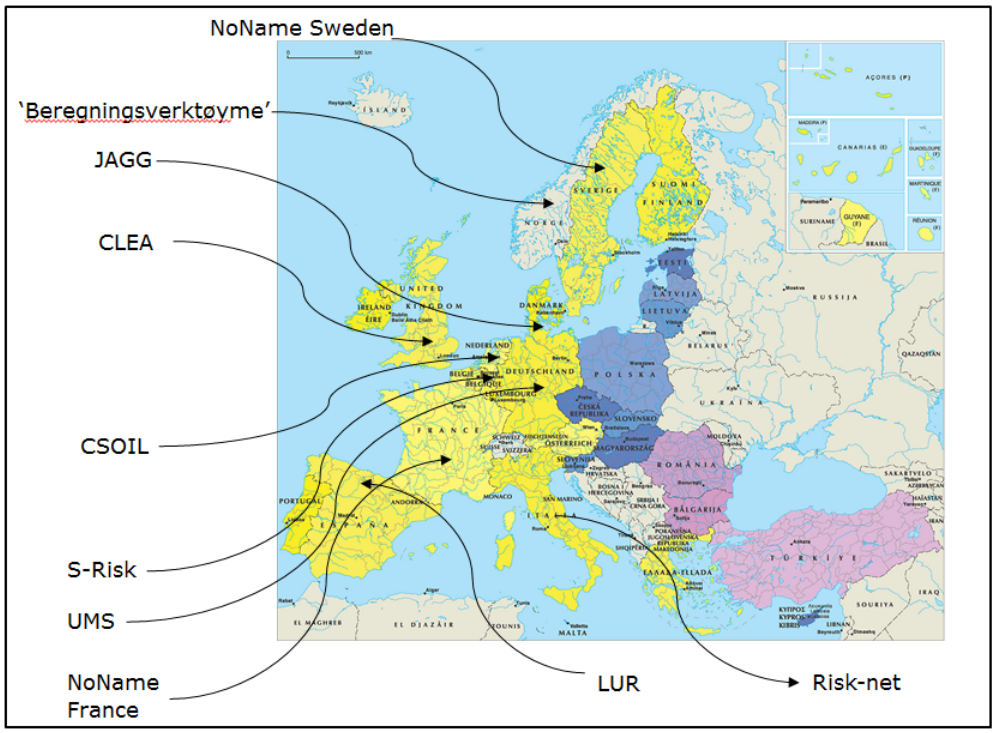


Figure 2. Flow chart of the Dutch CSOIL exposure model. Figure prepared by Frank Swartjes.



Moreover, procedures were developed for ecological risk assessment, including the Species Sensitivity Distributions (see section on SSDs), based on empirical relations between concentration in soil or groundwater and the percentage of species or ecological processes that experience adverse effects (PAF: potentially Affected Fraction). For site specific risk assessment, the TRIAD approach was developed, based on three lines of evidence, i.e., chemically-based, toxicity-based and using data from ecological field surveys (see section on the TRIAD approach).

In the framework of the HERACLES network, another attempt was made to summarizing different EU policies on polluted soil and groundwater. A strong plea was made for harmonisation of risk assessment tools (Swartjes et al., 2009). The authors also described a procedure for harmonization based on the development of a toolbox with standardized and flexible risk assessment tools. Flexible

tools are meant to cover national or regional differences in cultural, climatic and geological (e.g., soil type, depth of the groundwater table) conditions. It is generally acknowledged, however, that policy decisions should be taken on the national level. In 2007, an analysis of the differences of soil and groundwater screening values and of the underlying regulatory frameworks, human health and ecological risk assessment procedures (Carlon, 2007) was launched. Although screening values are difficult to compare, since frameworks and objectives of screening values differ significantly, a general conclusion can be drawn for e.g. the screening values at the potentially unacceptable risk level (often used as *'action' values*, i.e. values that trigger further research or intervention when exceeded). For the 20 metals, most soil screening values (from 13 countries or regions) show between a factor of 10 and 100 difference between the lowest and highest values. For the 23 organic pollutants considered, most soil screening values (from 15 countries or regions) differ by a factor of between 100 and 1000, but for some organic pollutants these screening values differ by more than four orders of magnitude. These conclusions are merely relevant from a policy viewpoint. Technically, these conclusions are less relevant, since, the screenings values are derived from a combination of different protection targets and tools and based on different policy decisions. Differences in screening values are explained by differences in geographical and biological and socio-cultural factors in different countries and regions, different national regulatory and policy decisions and variability in scientific/ technical tools.

References

Carlon, C. (Ed.) (2007). Derivation methods of soil screening values in Europe. A review and evaluation of national procedures towards harmonisation, JRC Scientific and Technical report EUR 22805 EN.

Elert, M., Bonnard, R., Jones, C., Schoof, R.A., Swartjes, F.A. (2011). Human Exposure Pathways. Chapter 11 in: Swartjes, F.A. (Ed.), Dealing with Contaminated Sites. From theory towards practical application. Springer Publishers, Dordrecht.

Swartjes, F.A. (2007). Insight into the variation in calculated human exposure to soil contaminants using seven different European models. Integrated Environmental Assessment and Management 3, 322-332.

Swartjes, F.A., D'Allesandro, M., Cornelis, Ch., Wcislo, E., Müller, D., Hazebrouck, B., Jones, C., Nathanail, C.P. (2009). Towards consistency in risk assessment tools for contaminated sites management in the EU. The HERACLES strategy from the end of 2009 onwards. National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands, RIVM Report 711701091.

Van den Berg, R. (1991/1994/1995). Exposure of humans to soil contamination. A quantitative and qualitative analyses towards proposals for human toxicological C-quality standards (revised version of the 1991/ 1994 reports). National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands, RIVM-report no. 725201011.

Further reading

Swartjes, F.A. (Ed.) (2011). *Dealing with Contaminated Sites. From theory towards practical application*. Springer Publishers, Dordrecht.

Rodríguez-Eugenio, N., McLaughlin, M., Pennock, D. (2018). *Soil Pollution: a hidden reality*. Rome, FAO.

6.5.6. Question 1

What is the logical first step in policy on soil and groundwater protection, related to ' prevention' ?

6.5.6. Question 2

What are the most frequently used protection targets on policies on soil and groundwater in the world?

6.5.6. Question 3

What role should screening values play in sophisticated risk assessment procedures?

6.5.6. Question 4

What is the ideal approach when developing a new soil and groundwater policy?

6.5.6. Question 5

Regarding the harmonization of risk assessment tools: why are flexible risk assessment tools necessary?

6.5.7. Drinking water

in preparation

This page titled 2.3.6.5: Regulatory Frameworks is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **6.5: Regulatory Frameworks** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.6.6: Risk management and risk communication

6.6. Risk management and risk communication

in preparation

This page titled 2.3.6.6: Risk management and risk communication is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **6.6: Risk management and risk communication** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.3.6.7: Risk perception

6.7. Risk perception

Author: Fred Woudenberg

Reviewers: Ortwin Renn

Learning objectives:

- To list and memorize the most important determinants of risk perception
- To look at and try to understand the influence of risks in situations or activities which you or others encounter or undertake in daily life
- To actively look for as many situations and activities as possible in which the annual risk of getting sick, being injured or to die has little influence on risk perception
- To look for examples in which experts (if possible ask them) react like lay people in their own daily lives

Key words: Risk perception, fear, worry, risk, context

Introduction

If risk perception had a first law like toxicology has with Paracelsus' "Sola dosis facit venenum" (see section on History of Environmental Toxicology) it would be:

"People fear things that do not make them sick and get sick from things they do not fear."

People can, for instance, worry extremely over a newly discovered soil pollution site in their neighborhood, which they hear about at a public meeting they have come to with their diesel car, and, when returning home, light an extra cigarette without thinking to relieve the stress.



Sources: Left: *File photo: Kalamazoo City Commissioner Don Cooney and Kalamazoo residents march to protest capping the Allied Paper Landfill, May 2013.* https://www.wmuk.org/post/will-bioremediation-work-allied-microbial-ecologist-skeptical. Right: https://desmotivaciones.es/352654/Obama

The explanation for this first law is quite easy. The annual risk of getting sick, being injured or to die has only limited influence on the perception of a risk. Other factors are more important. Figure 1 is a model of risk perception in its most basic form.

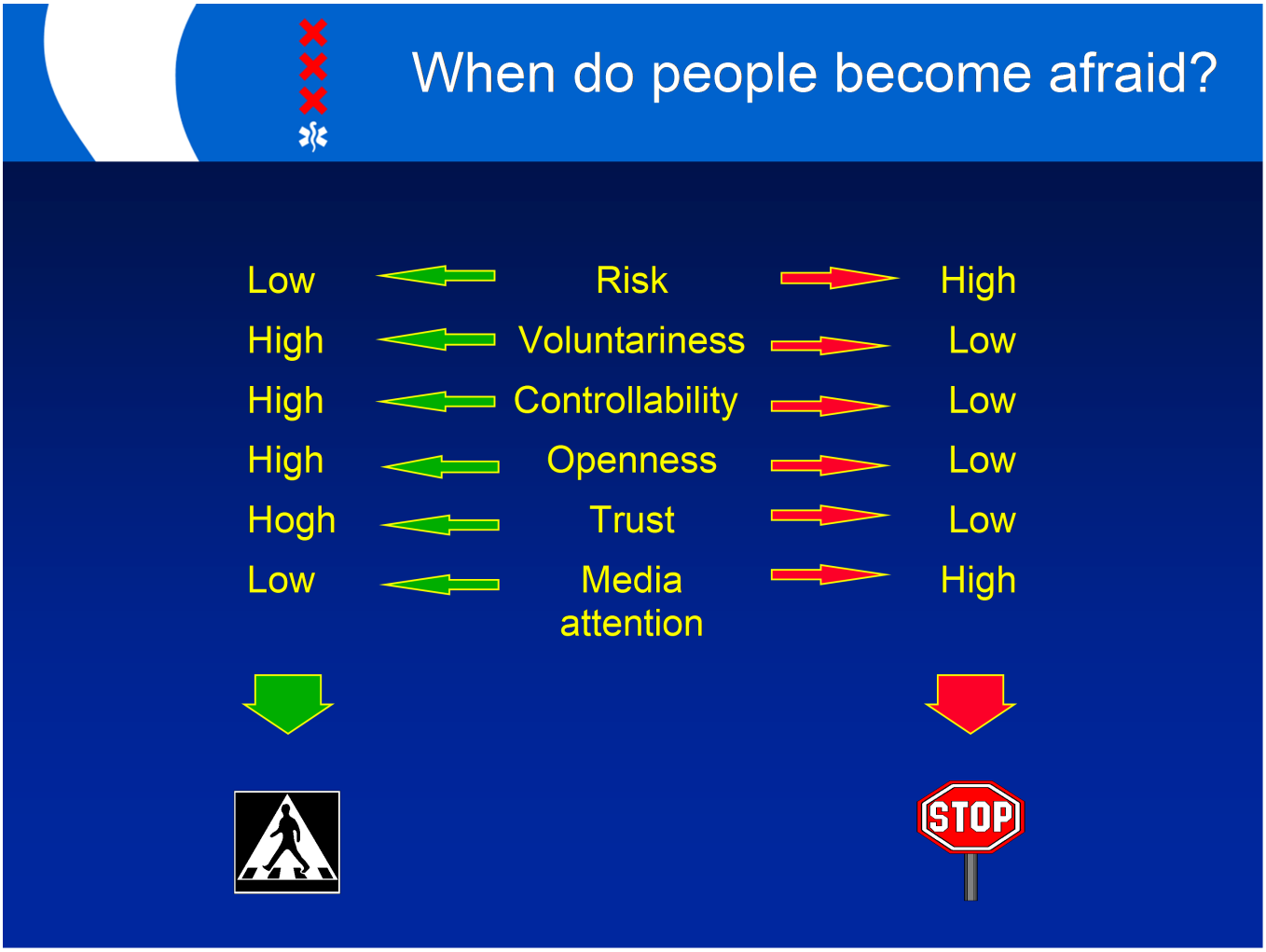


Figure 1. Simplified model of risk perception

In the middle of this figure, there is a list with several factors which determine risk perception to a large extent. In any given situation, they can each end up at the left, safe side or on the right, dangerous side. The model is a simplification. Research since the late sixties of the previous century has led to a collection of many more factors that often are connected to each other (for lectures of some well-known researchers see examples 1, 2, 3, 4 and 5).

Why do people fear soil pollution?

An example can show this interconnection and the discrepancy between the annual health risks (at the top of Figure 1) and other factors. The risk of harmful health effects for people living on polluted soil is often very small. The factor 'risk' thus ends at the left, safe side. Most of the other factors end up at the right. People do not voluntary choose to have a soil pollution in their garden. They have absolutely no control over the situation and an eventual sanitation. For this, they depend on authorities and companies. Nowadays, trust in authorities and companies is very small. Many people will suspect that these authorities care more about their money than about the health and well-being of their citizens and neighbours. A newly discovered soil pollution will get local media attention and this will certainly be the case if there is controversy. If the untrusted authorities share their conclusion that the risks are low, many people will suspect that they withhold information and are not completely open. Especially saying that there is 'no cause for alarm' will only make people worry more (see a funny example). People will not believe the conclusion of authorities that the risk is low, so effectively all factors end up at the dangerous side.

Why smokers are not afraid

For smoking a cigarette, the evaluation is the other way around. Almost everybody knows that smoking is dangerous, but people light their cigarette themselves. Most people at least think they have control over their smoking habit as they can decide to stop at

any moment (but being addicted, they probably highly overestimate their level of control). For their information or taking measures, people do not depend on others and no information is withheld about smoking. Some smokers suffer from what is called optimistic bias, the idea that misery only happens to other. They always have the example of their grandfather who started smoking at 12 and still ran the marathon at 85.

People can be upset if they learn that cigarette companies purposely make cigarettes more addictive. It makes them feel the company takes over control which people greatly resent. This, and not the health effects, can make people decide to quit smoking. This also explains why passive smoking is more effective than active smoking in influencing people's perceptions. Although the risk of passive smoking is 100 times smaller than the risk of active smoking, most factors end up at the right, dangerous side, making passive smoking maybe 100 times more objectionable and worrisome than active smoking.

Experts react like lay people at home

Many people are surprised to find out that the calculated or estimated health risks influences risk perception so little. But we experience it in our own daily lives, especially when we add another factor in the model: advantages. All of us perform risky activities because they are necessary, come with advantages or sometimes out of sheer fun. Most of us take part in daily traffic with an annual risk of dying far higher than 1 in a million. Once, twice or even more times a year we go on a holiday with a multitude of risks: transport, microbes, robbery, divorce. The thrill seekers of us go diving, mountain climbing, parachute jumping without even knowing the annual fatality rates. If the stakes are high, people can knowingly risk their life in order to improve it, as the thousands of immigrants trying to cross the Mediterranean illustrate, or even to give their life for a higher cause, like soldiers at war (Winston Churchill in 1940: "I have nothing to offer but blood, toil, tears and sweat").

An example at the other side can show it maybe even clearer. No matter how small the risk, it can be totally unacceptable and nonsensical. Suppose the government starts a new lottery with an extremely small chance of winning, say one in a billion. Every citizen must play and tickets are free. So far nothing strange, but there is a twitch. The main and only price of the lottery is a public execution broadcasted live on national TV. The government will probably not make itself very popular with this absurd lottery. When government, as still is done, tells people they have to accept a small risk because they accept larger risks of activities they choose themselves, it makes people feel they have been given a ticket in the above mentioned lottery. This is how people can feel if the government tells them the risk of the polluted soil they live on is extremely small and that it would be wiser for them to quit smoking.

All risks have a context

A main lesson which can be learned from the study of risk perception is that risks always occur in a context. A risk is always part of a situation or activity which has many more characteristics than only the chance of getting sick, being injured or to die. We do not judge risks, we judge situations and activities of which the risk is often only a small part. Risk perception occurs in a rich environment. After 50 years of research a lot has been discovered, but predicting how angry of afraid people will be in a new, unknown situation is still a daunting task.

6.7. Question 1

Name at least 5 important determinants of risk perception.

6.7. Question 2

Do people judge risks in isolation or the situation or activities they are part of?

6.7. Question 3

How large in general is the influence of the risk on risk perception?

6.7. Question 4

Do experts react differently from lay people if they encounter risks in their own daily lives?

This page titled 2.3.6.7: Risk perception is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Sylvia Moes, Kees van Gestel, & Gerco van Beek via source content that was edited to the style and standards of the LibreTexts platform.

- **6.7: Risk perception** by Sylvia Moes, Kees van Gestel, & Gerco van Beek is licensed CC BY 4.0. Original source: https://maken.wikiwijs.nl/147644/.

2.4: Key Elements of Green Chemistry (Lucia)

Green chemistry, in addition to being a science, it is also a philosophy and nearly a religion. Attendance at American Chemical Society Green Chemistry & Engineering Conferences will instill such an ideal into any attendant because of the nearly universal appeal and possibilities in this novel approach to radicalizing the business of doing science and engineering.

This page titled 2.4: Key Elements of Green Chemistry (Lucia) is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

Front Matter

Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

SECTION OVERVIEW

2.4.1: Key Elements of Green Chemistry

2.4.1.1: Introduction to Green Chemistry

2.4.1.2: Green Chemistry Concepts

2.4.1.3: Case Study

2.4.1.4: Conclusions and Review

This page titled 2.4.1: Key Elements of Green Chemistry is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

2.4.1.1: Introduction to Green Chemistry

Introduction

The rise of the industrial complex and society's strong need to increase its quality of life, security, wealth, and future productivity have helped to relegate the discipline of chemistry and its work to a very lofty position of prominence. Our progress toward achieving a Utopia of human intellectual growth, freedom of choices, equality within classes, and wealth acquisition have been supplemented and encouraged by the fruits of chemical research & development efforts since the dawn of last century. For example, we have witnessed significant strides in the following facets of our lives:

- Water: the introduction of reagents such as chlorine and fluoride to not only reduce and eliminate microbial contamination, but to also improve dental health, respectively;
- Food: we have seen the ability to use fluoro, chlorocarbons (CFCs, "Freon") to help refrigerate foods for much longer "shelf life", packaging that has specific mechanical and antimicrobial properties, and preservatives/food processing aids for eliminating spoilage;
- Clothing: synthetic fibers such as rayon, nylon, polyester, and Kevlar have been indispensable to not only normal wear, but to military, geologic, agricultural, and packaging applications;
- Safety: materials chemistry has spurred the manufacture of lightweight, but tough materials such as helmets, resins, plastics, etc., that have helped save lives in a number of military and industrial sectors.

The chemical industry has played a vital role in the emergence of high quality life within our human societies in these latter areas as well as many others. In fact, the industry is among the top ten industrial sectors in terms of gross output and sales. One of the most adverse realities and perceptions of the industry is its impact on the environment. Although it is a very safe industry as a whole, well publicized disasters have contributed to a poor public perception. For example, issues such as eutrophication, persistent organic pollutants, BOD, and the famous burning Cuyahoga River are typical of the calamities that damage the reputation and importance of our industry.

Please see http://clevelandhistorical.org/items/show/63#.V7NI5BWU2og

In fact, the morning of June 22, 1969 was witness to a fire from the burning of oil and debris that collected on the surface of the Cuyahoga River (Cleveland). The event roused up a media storm, captured national attention, and was featured in a commentary on the nation's environmental problems (Time magazine, Aug. 1). It was a clarion call to how calamitous the nation's environmental problems had become. The then EPA Administrator Lisa Jackson commented years later (2011) that the fire evidence of "the almost unimaginable health and environmental threats" from water pollution of the time. Clearly, as one environmentalist stated, "when rivers are on fire, you know things are bad." The image was seared into the nation's emerging environmental consciousness and fueled a demand for greater regulation. In 1972, Congress passed the federal Clean Water Act (CWA). Today, the nation's waters are far cleaner, and many credit CWA with preventing other rivers from befalling a similar fate.

The reality is that the fire was a symbol of how bad river conditions in the US had once been. That fire was not the first time an industrial river had caught on fire, but it was the last. For example, the late 19th and early 20th century saw many river fires. At least 13 on the Cuyahoga occurred, whereas there were others on rivers in Detroit, Baltimore, Philadelphia, and elsewhere. Although industries, chemists, and governments do not ever intentionally wish to or contribute to cause harm, the law espoused by "Murphy" holds. The contemporary form of the law can be traced back to 1952 as an epigraph for a mountaineering book by Sack who described the adage as an "ancient mountaineering adage":

Anything that can possibly go wrong, does (Sack, 1952).

Thus, the challenge for contemporary chemists and allied workers in view of the law is to ensure that any and all reactions, processes, and designs keep this general philosophy in mind. Our goal in the twenty first century is to mitigate potential disasters that can occur from our chemistries and engineering.

Sustainable Development

The concept of "sustainable development" was articulated from a UN Commission on Environment and Development in 1987 (Brundtland Commission). It is simply stated:

'..meeting the needs of the present without compromising the ability of future generations to meet their own needs.' Since the Brundtland Commission, a number of governments, NGOs, and societies have considered deeply what sustainable really means. Several of the key drivers heating up this
discussion include:

1. What is an acceptable rate of depletion of fossil fuels? Is there?
2. Is there an acceptable level of pollution to release into the atmosphere and water?

Obviously, waste is a given in any industrial process, but the primary question is not its generation, but how it is handled. If we produce a waste stream, can it be recovered, recycled, or reused? If carbon dioxide is the waste stream, can we conclude that nature will sequester it or will it contribute to global warming (climate change)? These are questions that warrant intense analysis and follow-up. The overall fate of the planet (humanity, ecosystems, and our way of life) hang in the balance. Methods to address ensuring the continuation of our planet and its resources include:

- Not allowing for the accretion of toxics (e.g., heavy metals) from the earth's crust and other
- beds;
- Not continuously creating non-degradable/perishable compounds (e.g., CFC's, benzodioxocins) that can cause significant damage to the ozone and aquatic life, respectively;
- Ensuring that the natural processes in place on earth are not disrupted (e.g., ravaging
- rainforests, polluting watersheds);
- Not depleting or hoarding natural resources (e.g., water).

We as an educated community (students, teachers, government workers, industry workers, etc.)recognize and appreciate that the Earth is equipped to deal with hiccups and disturbances in its processes, but continuously pushing such boundaries of recovery will only lead to disasters.Indeed, a paradigm for ensuring that our chemical processes do not exceed the capacity of the Earth to engage in buffering and recovery is the following:

To promote innovative chemical technologies that reduce or eliminate the use or generation of hazardous substances in the design, manufacture, and use of chemical products.

The above statement defines the concept of "Green Chemistry." The subject has now become a very well accepted and welcomed part of many chemistry curricula and industry philosophies. In fact, the concept of a triple bottom line, i.e., financial, social, and environmental owes its existence to green chemistry. Green chemistry has become an incredibly indispensable field within the panoply of chemistry subject matter. Its content and mission, however, are very unique in comparison to its sister courses because it embodies a social and environmental focus.

This page titled 2.4.1.1: Introduction to Green Chemistry is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **1.1: Introduction to Green Chemistry** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.1.2: Green Chemistry Concepts

Agricultural Support

The growing of crops in different farming regions to feed the masses owes its potential greatly to harnessing the power of chemicals such as herbicides, pesticides, fungicides, etc. Although in general we tend to think of the benefits of pesticides, there are a number of issues associated with their use that go well beyond their ability to promote the cultivation and availability of food. For example, DDT (Figure 2.4.1.2.1) is a chemical that has seen much use in the US. Over the period of 1950 to 1980, it was used in agriculture at the rate of more than 40,000 tons each year worldwide and it has been estimated that a total of 1.8 million tons have been produced around the world since the 1940s. In the United States, it was manufactured by numerous companies including Monsanto, Ciba, Montrose Chemical Company, and Pennwalt. More than 600,000 tons (1.35 billion pounds) were applied in the US before it was banned in 1972.

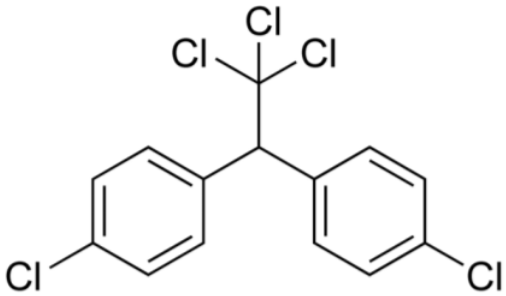


Figure 2.4.1.2.1: Shown above is a chemical representation of DDT or DicholoroDiphenylTrichloroethane, a colorless, crystalline, tasteless and nearly odorless organochlorine compound that displays insecticidal properties with huge environmental impacts. DDT has been formulated in multiple forms that include solutions in xylene or petroleum distillates, emulsifiable concentrates, water-wettable powders, including solutions in vaporizers and lotions. https://www.wikiwand.com/en/Dichlorodiphenyltrichloroethane

Nature's Best

The idea that nature produces chemicals that are "green" is a fallacy or a myth. The concept of green deserves some clarification at this point. What is green and what isn't is actually a matter of nature, quantity, human safety, long term effects, and acute toxicity. Green is generally a term associated with a sustainable (renewable) product or non-toxic process whose employment in society has no acute toxicity and a general favorable life cycle analysis. There are a number of documented chemicals in nature that are extremely toxic even at small doses thus invalidating the idea that nature is "green". For example, aflatoxin B1, shown below in Figure 2.4.1.2.2, is a toxin produced by *Apsergillus flavus* and *A. parasiticus* that is one of the most potent carcinogens known. It is a common contaminant of a variety of foods including peanuts, cottonseed meal, corn, and other grains as well as animal feeds. According to the Food and Agriculture Organization, the worldwide maximum levels of aflatoxin B1 is in the range of 1–20 µg/kg in food, and 5–50 µg/kg in dietary cattle feed.

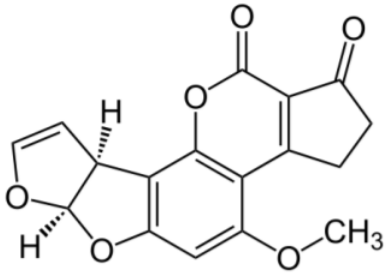


Figure 2.4.1.2.2: Shown above is a chemical representation of Aflatoxin B1, one of the most notorious carcinogens available. And it is "natural"! The IUPAC name is: 6aR,9aS)-2, 3, 6a, 9a-Tetrahydro4-methoxy-1H, 11H cyclopenta[c]furo[3',2':4,5]furo[2,3- h][1]benzopyran-1,11-dione. https://www.wikiwand.com/en/Aflatoxin_B1

Shown in Figure 2.4.1.2.3 below is a derivative of lysergic acid. Lysergic acid, also known as D-lysergic acid and (+)-lysergic acid, is a precursor to a diverse array of ergoline alkaloids produced by the ergot fungus and found in the seeds of *Turbina corymbosa* (ololiuhqui), *Argyreia nervosa* (Hawaiian Baby Woodrose), and *Ipomoea tricolor* (morning glories, tlitliltzin). Amides of lysergic acid, lysergamides (see Fig. 3), are widely used as pharmaceuticals and as psychedelic drugs (LSD). Lysergic acid received its name from the lysis of various ergot alkaloids.

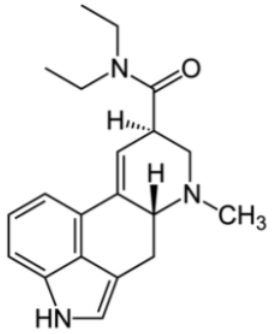


Figure 2.4.1.2.2*:* Shown above is the diethylamide version of lysergic acid.
https://www.wikiwand.com/en/Lysergic_acid

Again, as already stipulated, nature's best does not necessarily imply that her products are "green". Green, by virtue of the product distribution found in nature (e.g. cocaine), can certainly pose serious if not lethal implications for human consumption. We all know that sodium chloride is ubiquitous within nature. It is a common salt that manifests a cubic crystal lattice structure that readily dissolves in water where it is mostly found. Humans consume it in great quantities as a flavorant, seasoning, and preservative. Nevertheless, its mass consumption especially in the US has led to > 65 million people afflicted with high blood pressure while numerous others have all other maladies associated from its overconsumption.

Accidents & Waste Minimization

Waste because of the nature of living processes is a natural consequence. Within the schema of living cell frameworks, the ability to generate waste is considered a clear indication of a "living" system. Waste is a necessary by-product of any work function even if it is heat, light, or some other form of energy by virtue of the intrinsic inefficiencies associated with the utilization of "fuel" or raw materials. For example, even a newborn that is fed on breast milk, a ideal natural food source for babies, does not simply produce a non-material, i.e., heat/energy dissipation, waste product. The problems associated with waste including inherent inefficiencies with using resources are how it builds up and affects the quality of the environment. In general, we cannot prevent waste because everything is subject to entropy.

The crux of the problem associated with waste boils down to safety. As alluded to in the Prezi example above, rust can cause material fatigue or loss of integrity that can compromise functionality. Additionally, chemical processes that produce waste need to adhere to the three general "R"s of Green Chemistry.

the Three "R"s of Green Chemistry

- Reduce
- Reuse
- Recycle

This page titled 2.4.1.2: Green Chemistry Concepts is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **1.2: Green Chemistry Concepts** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.1.3: Case Study

The LeBlanc Case

The Leblanc process industrial factories were very damaging to the local environment. It was an early industrial process for producing *soda ash* (sodium carbonate) that was named after its inventor. It required two stages: sodium sulfate (salt cake) produced from sodium chloride (salt, we just discussed this!) followed by reaction with coal and calcium carbonate to produce sodium carbonate. It eventually became obsolete after development of the Solvay process. The process of generating salt cake involved reacting salt with sulfuric acid which released hydrochloric acid gas, an acid that was industrially useless in the early 19th century and therefore vented into the atmosphere. Also, an insoluble, smelly solid waste (CaS, *galligu*) was produced. The inefficiency of the process was horrific! Each ton of *soda ash*, the process produced ~1.5 tons of hydrogen chloride and >1 ton of calcium sulfide, the useless waste product! *Galligu* had no economic value so it was piled in heaps and spread on fields where it weathered to release
hydrogen sulfide (what a smell!).

Leblanc soda works then became targets of lawsuits and legislation. A lawsuit from 1839 alleged that, “the gas from these manufactories is of such a deleterious nature as to blight everything within its influence and is alike baneful to health and property. The herbage of the fields in their vicinity is scorched, the gardens neither yield fruit nor vegetables; many flourishing trees have lately become rotten naked sticks. Cattle and poultry droop and pine away. It tarnishes the furniture in our houses, and when we are exposed to it, which is of frequent occurrence, we are afflicted with coughs and pains in the head ... all of which we attribute to the Alkali works.”

Therefore, in 1863, the British Parliament passed the first Alkali Act, a precursor to the first modern air pollution legislation. This Act dictated that no more than 5% of the hydrochloric acid produced by alkali plants could be vented. To comply, soda works passed it up a tower packed with charcoal where it was absorbed by water flowing in the other direction. Unfortunately, the chemical works usually dumped the resulting solution into nearby bodies of water where it promptly ended up killing fish and other aquatic life.

This page titled 2.4.1.3: Case Study is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **1.3: Case Study** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.1.4: Conclusions and Review

The greatest blight affecting the modern time is the apathy associated with ignorance. We tend to strongly adhere to wrongful beliefs—all manufacturing and industrial processes are failsafe in protecting humans and the environment. The truth is that Murphy's Law more often than not is RIGHT. Thus, the need for pre-emptive measures as embodied by the principles of green chemistry is of paramount significance. The principles as listed below will constitute a future chapter in this book:

Prevention
It is better to prevent waste than to treat or clean up waste after it has been created.

Atom Economy
Synthetic methods should be designed to maximize the incorporation of all materials used in the process into the final product.

Less Hazardous Chemical Syntheses
Wherever practicable, synthetic methods should be designed to use and generate substances that possess little or no toxicity to human health and the environment.

Designing Safer Chemicals
Chemical products should be designed to affect their desired function while minimizing their toxicity.

Safer Solvents and Auxiliaries
The use of auxiliary substances (e.g., solvents, separation agents, etc.) should be made unnecessary wherever possible and innocuous when used.

Design for Energy Efficiency
Energy requirements of chemical processes should be recognized for their environmental and economic impacts and should be minimized. If possible, synthetic methods should be conducted at ambient temperature and pressure.

Use of Renewable Feedstocks
A raw material or feedstock should be renewable rather than depleting whenever technically and economically practicable.

Reduce Derivatives
Unnecessary derivatization (use of blocking groups, protection/ deprotection, temporary modification of physical/chemical processes) should be minimized or avoided if possible, because such steps require additional reagents and can generate waste.

Catalysis
Catalytic reagents (as selective as possible) are superior to stoichiometric reagents.

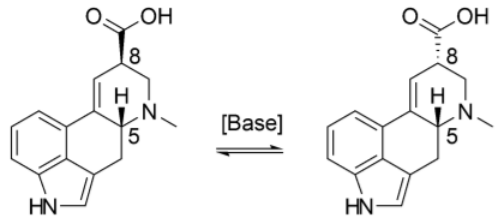
Design for Degradation
Chemical products should be designed so that at the end of their function they break down into innocuous degradation products and do not persist in the environment.

Real-time analysis for Pollution Prevention
Analytical methodologies need to be further developed to allow for real-time, in- process monitoring and control prior to the formation of hazardous substances.

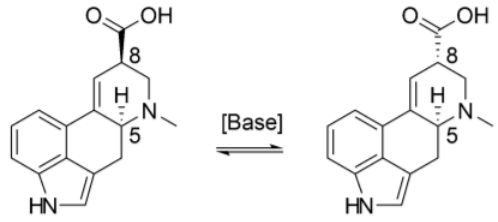
Inherently Safer Chemistry for Accident Prevention
Substances and the form of a substance used in a chemical process should be chosen to minimize the potential for chemical accidents, including releases, explosions, and fires.

Review Questions

2.4.1.4.1 Using lysergic acid as a starting point, indicate the general type of chemistry (you can indicate in your own words what is happening) that would need to occur in order for the epimerization of lysergic and L-lysergic acid to occur shown below. https://www.wikiwand.com/en/Lysergic_acid



Lysergic acid **Isolysergic acid**



L-Lysergic acid **L-Isolysergic acid**

2.4.1.4.2.How does too much salt in a human's diet contribute to high blood pressure?

a. What are two known chemical mechanisms (reactions) that enzymes can pursue for bonding to 4-hydroxy-2-nonenal?
b. Come up with a way to quantify the inefficiency of a reaction; that is, derive a quantitative measure of the wastefulness of a reaction using the LeBlanc numerical process information contained in this chapter.

This page titled 2.4.1.4: Conclusions and Review is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **1.4: Conclusions and Review** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

SECTION OVERVIEW

2.4.2: Life-Cycle Analysis

2.4.2.1: Introduction to Life Cycles

2.4.2.2: LCA/LCIA Concepts

2.4.2.3: Conclusions and Review

This page titled 2.4.2: Life-Cycle Analysis is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

2.4.2.1: Introduction to Life Cycles

Green chemistry, in addition to being a science, it is also a philosophy and nearly a religion. Attendance at American Chemical Society Green Chemistry & Engineering Conferences will instill such an ideal into any attendant because of the universal appeal and possibilities in this novel approach to radicalizing the business of doing science and engineering. Life Cycle Assessment (LCA) is a comprehensive life cycle approach that quantifies ecological and human health impacts of a product or system over its complete life cycle. It uses credible scientific methods to model steady-state, global environmental and human health impacts. It also helps decision-makers understand the scale of many environmental and human health impacts of competing products, services, policies or actions.



Please see for more information https://www.youtube.com/watch?v=NQTW7jjXVmE (LCA: Intro to Life Cycle Assessment. (2013).)

Triple bottom line (economic, social, and environmental)

Triple Bottom Line accounting enables enterprises to support sustainability analysis in their operations, products and services. LCA contributes to the Triple Bottom Line reporting by quantifying the ecological and human health performance of competing products and services (Figure 2-1). Adding the social and economic performance reporting of a product or service to the LCA results of the product or service is one way to deliver Triple Bottom Line reporting.

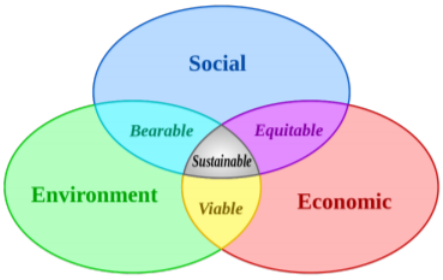


Figure 2.4.2.1.1: The triple bottom line in GC containing social, economic, and environmental Venn diagram whose intersection contributes to a valid sustainability analysis.

LCA for Decision Making

Who makes decisions

- Company product managers or planners
- Company procurement and purchasing
- Industrial sector consortia (example: aluminum manufacturers)
- Regional or national policy makers
- Consumers, customers and product users

Primary drivers and expectations for LCA:

- Learning about the environmental performance of products and services

- Minimizing production and regulatory costs
- Minimizing environmental and human health damage
- Understanding trade-offs between multiple impact categories and product phases
- Supporting equitable economic distribution and profitable operations

Types of decision situations

- strategic planning and capital investments
- (green building, waste management)
- eco-design, product development
- operational management
- (green procurement)
- communication and marketing
- (eco-labeling, product information)

Eco-labels:

- Type 1: Third-party certified multi-criteria environmental labeling. Example: Forest Stewardship Council label for wood products
- Type 2: Environmental self-declaration claims. Example: Single issue claim such as the "bio-degradable" label
- Type 3: Independently verified label with preset quantified indices. Example: Numerical water consumption rating for a dish washer

Eco-labels that require LCA:

Type 1) Third-party certified multi-criteria environmental labeling

Type 3) Independently verified label with preset quantified indices

Life Cycle Assessment (LCA) enables the creation of Type 1 and Type 3 eco-labels. These eco-labels can be powerful tools in obtaining larger shares of a specific market sector. Life cycle assessment (LCA) is a standardized programmatic tool to determine the environmental impacts of products or services. It can be described by a four-part framework as outlined by the 14044 ISO standard.

Life cycle assessment Framework

1. Goal and scope definition
2. Life cycle inventory
3. Life cycle impact assessment
4. Interpretation

This integrated framework was inspired by earlier forms of life-cycle thought from life cycle financial analysis. Examining a product from origination to use and disposal provides more holistic analysis to identify where environmental impacts originate and guide efforts in reducing impacts.

The ISO standards (http://www.iso.org/iso/home/standards.htm) provides guidance on structure framework, reuse requirements of data, study assumptions, and methods. By using more standard LCA methodologies, studies are more comparable and of greater scientific rigor. A standardized method allows LCA practitioners to manage complex datasets, provide comparisons among products, and allow benchmarking. In the absence of a standardized method, the results of LCA studies are even more variable depending on study assumptions and methods. The ISO standards help reduce the influence of the practitioner influence on study results.

This page titled 2.4.2.1: Introduction to Life Cycles is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **2.1: Introduction to Life Cycles** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.2.2: LCA/LCIA Concepts

Concepts

Goal, Scope, and Definition

Assumptions inherent in an LCA study are apt to change the results and conclusions derived from analysis. In addition, many different types of studies require various levels data collection and analysis. The goal and scope of a LCA defines its intent, targeted audience, and use. The intended use informs further decisions for scope, functional unit of comparison, and data collection. For example, if a LCA study is used internally, a full review panel of LCA experts is not required; however, when providing public environmental claims about a competing product, a review is required.

Inventory analysis

A life cycle inventory (LCI) is the most laborious step of a LCA: data is collected and organized. It often involves contacting companies, accumulating literature sources, and building models using life cycle assessment software. Materials flows, types of materials, product life time, and product energy requirements are collected in the LCI phase.

Life Cycle Impact Assessment

A life cycle impact assessment (LCIA) part of the analysis process collects life cycle inventory data and delivers environmental impacts values. This process greatly reduces the complexity of the data set from hundreds of inputs to 10 or fewer impact categories for decision-making. There are many different methods for LCIA based on location, goals, and scope.

Interpretation

The interpretation step derives from what was found in the other steps for the generation of new information. It is not the last step but iterative. When it is done, the study assumptions, goals, scope, and methods are refined to suit the needs of the study.

Life Cycle Analysis: Goal, Scope and Boundaries

Goal

The first step is defining the goal to give the aim and what it encompasses. There are two types of LCA objectives: (1) descriptive and (2) change-oriented. The descriptive types look at broader aspects of an issues, e.g., how much of the world's carbon dioxide emissions are derived from commuters (light duty vehicles). These broader environmental questions fall within the domain of descriptive LCAs. The second type of LCA is change-oriented, in which two options for fulfilling a function are compared. Typical examples of change-oriented LCAs are paper vs. plastic, flying vs. driving, and gas vs. electric heating. These types of studies can guide the choice of methods to reduce environmental impacts. The intended audience is another part of the goal and scope. The audience may include interest groups such as policy makers, company marketing groups, or product development teams. Additionally, interest groups should be identified. These include companies, funding sources, target audiences, and expert reviewers. It is noted that the intended use of the LCA may be different from the end use because the information may be relevant to other decisions and analyses beyond the original intent.

One specific LCA to compare two products is a "comparative assertion disclosed to the public". In this type of study, "environmental claims regarding the superiority or equivalence of one product vs a competing product which performs the same function" are communicated. These types of studies must follow ISO 14044 standards with the nine steps for a "comparative assertion".

Scope

The scope definition serves the purpose of communicating to the audience what is included and what is excluded. Depending on the goal, there are several types of scopes including cradle-to-gate, cradle-to-grave, and gate-to-gate. There are other words commonly used to describe these scopes:

- **Cradle-to-grave:** includes all flows and impacts from obtaining raw material to disposal and reuse
- **Crade-to-gate:** includes all flows and impacts from raw material to production, but excludes product use and end of life.
- **Gate-to-gate:** only includes flows from production or material processing steps of a product life.

The scope must be carefully selected in consideration of the potential implications not including product stages or phases in the scope of the work. For example, a product may have lower product emissions, but have a shorter lifetime than an alternative product that would not be communicated in a process stage diagram as seen in Figure 2.4.2.2.1**.** These types of diagrams list the major unit steps considered and clearly show what is not included.

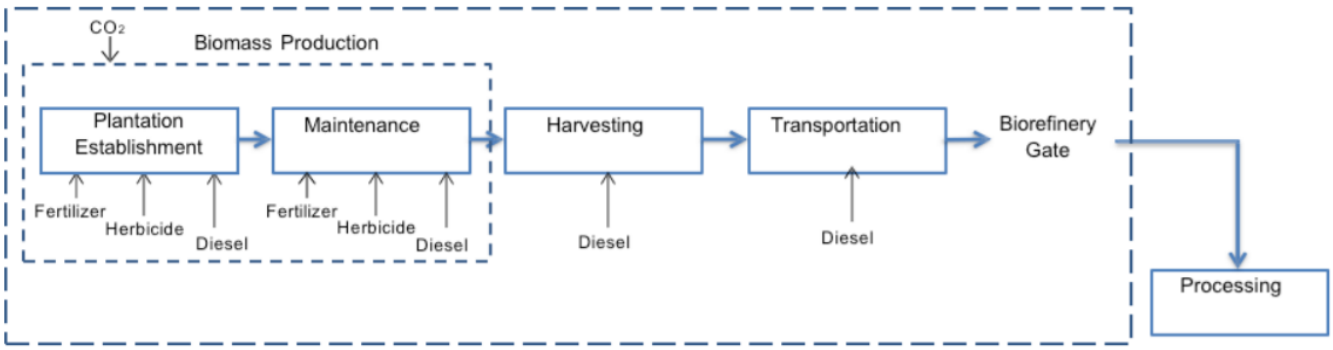


Figure 2.4.2.2.1: System boundary diagram of a cradle-to-grave biofuels process.

Temporal boundaries are also established in the scope. Assumptions relating to time can have a large influence on the results. A study timeframe should be picked which will best capture the impacts of the product or processes. A 100-year window is a common temporal boundary, for example, in global warming. In a 100-year temporal window, impacts occurring after 100 years are not part of the overall analysis.

Other aspects to be included in scope are technology and geographical regions. Many studies are spatially dependent so LCA results are not broadly applicable to other regions. Products or services from older technologies often have different impacts than current technologies. Thus, it is important to communicate the type and stage of the technology. In addition, allocation procedure impact assessment methods and should be reported.

Functional Unit

A functional unit is the primary measure of a product or service. ISO states that “the functional unit defines the quantification of the identified functions (performance characteristics) of the product. The primary purpose of a functional unit is to provide a reference to which the inputs and outputs are related. This reference is necessary to ensure comparability of LCA results. The functional unit can be a service, mass of material, or an amount of energy. Selecting appropriate functional units is critical to creating an unbiased analysis. For example, when comparing trains to cars for transportation, the comparison may suffer from the inability to correlate energy inputs and outputs. The real purpose of the train would be to deliver a larger number of people to a specific centralized location. For this example, a better functional unit may be impacts of a train delivering a specific number of people over a specified distance. The results will then be normalized to distance for more reasonable correlations and assessments.

Cut of Criteria

Data collection for an LCA is the most time-intensive and laborious step. Cut-off criteria are used to expedite the process. Cut-off criteria define a level of product content or other parameter to which the study will not consider. One example: materials contents less than 1% of the total product mass are not considered. This allows the LCA practitioner to focus on data from the main flows of the system while systematically eliminating flows which may not influence the results.

Life Cycle Inventory

In addition to data collection, the life cycle inventory (LCI) step is a very laborious aspect of life cycle assessment. The data collected for the product, production process, and product life cycle are used in the impact assessment to determine the environmental impacts. Collecting consistent, transparent, and accurate LCI data is critical to the success of an overall LCA.

Primary and Secondary Data

Example data collected:

- Raw material use
- Energy use

- Transportation distances
- Chemical use
- Waste treatment information
- Process yields
- Life time
- Water use
- Product and co-product flows
- Other flows in or out of the system that are within the defined cut-of-criteria

Tracking the material flows into and out of the defined system is the first step of LCI. After the materials flows have been determined through interviews, literature searches, and measurement, LCA software can be used to track the material process's elementary flows to and from the environment. Elementary flows originate in the environment and are mined or retrieved for use in a process or flows that are released from processes to the environment and are not used by other processes. These elementary flows are the actual materials used and materials released to the environment as a result of the studied product system. In Figure 2.4.2.2.2, the two types of LCI data can be seen. On the top half of the figure, process flows such as products, services and other goods are listed. The lower half lists elementary flows such as chemicals released to soil or air.

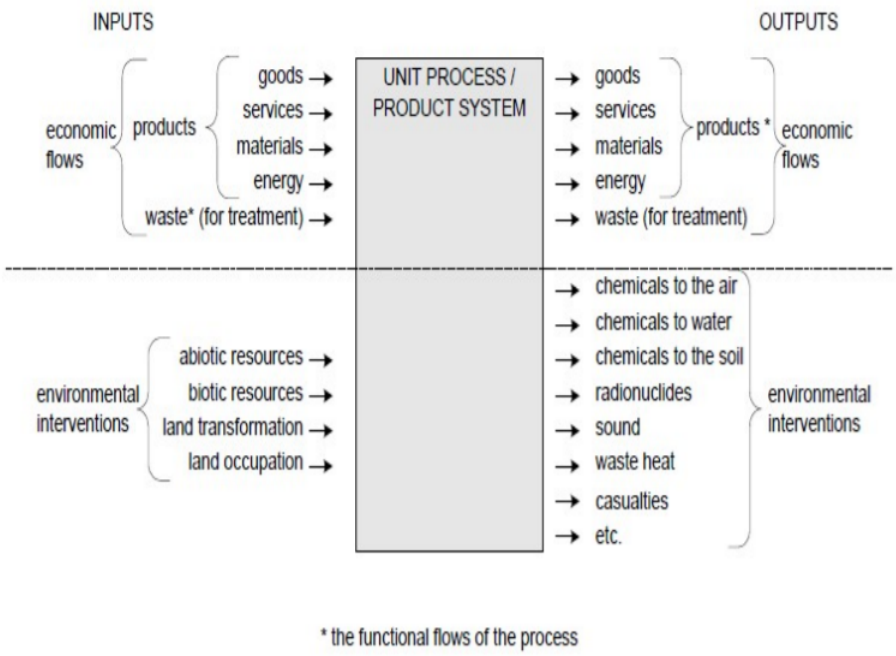


Figure 2.4.2.2.2: System boundary diagram of a cradle-to-grave biofuels process.

Life Cycle Impact Assessment

Life cycle impact assessment (LCIA) is among the last steps of LCA. The purpose of a LCIA "is to provide additional information to assess life cycle inventory (LCI) results and help users better understand the environmental significance of natural resource use and environmental releases". The LCIA helps provide significance and results for easier decision making; however, it is important to understand it does not directly measure the impacts of chemical releases to the environment as an environmental risk assessment does. The third step of LCIA follows sequentially after the LCI using the many flows to and from the environment developed in the LCI. These LCI flows, without an impact assessment step, are not easily interpreted and understanding the significance of emissions is impossible. The LCIA is different from a risk assessment measuring absolute values of environmental impacts in that the LCIA helps determine the significance of emissions and impacts in relation to the study scope. The absolute value of the impacts cannot be determined by the LCIA due to (Margni and Curran 2012):

- The relative expression of potential environmental impacts to a reference unit
- The integration of environmental data over space and time
- The inherent uncertainty in modeling environmental impact

- The fat that some possible environmental impact occurs in the future

Even though the LCIA has limitations, it is useful in determining what impacts matter, what unit processes are contributing the most through hot spot analysis and identify best scenario options when environmental tradeoffs occur.

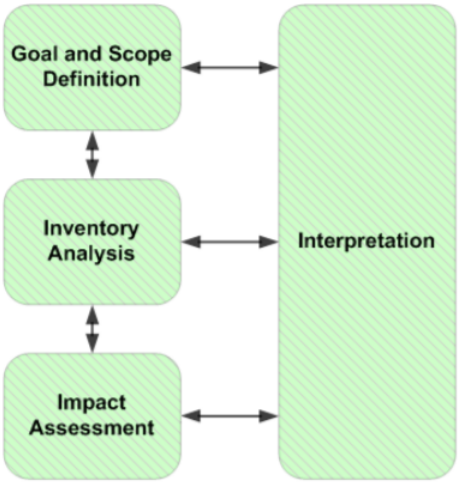


Figure 2.4.2.2.3: Life cycle assessment stages.
https://commons.wikimedia.org/w/index.php?curid=40862556

According to ISO there are three mandatory processes of a LCIA including Selection of impact categories, Classification, and Characterization, Figure \(\PageIndex{4}\).

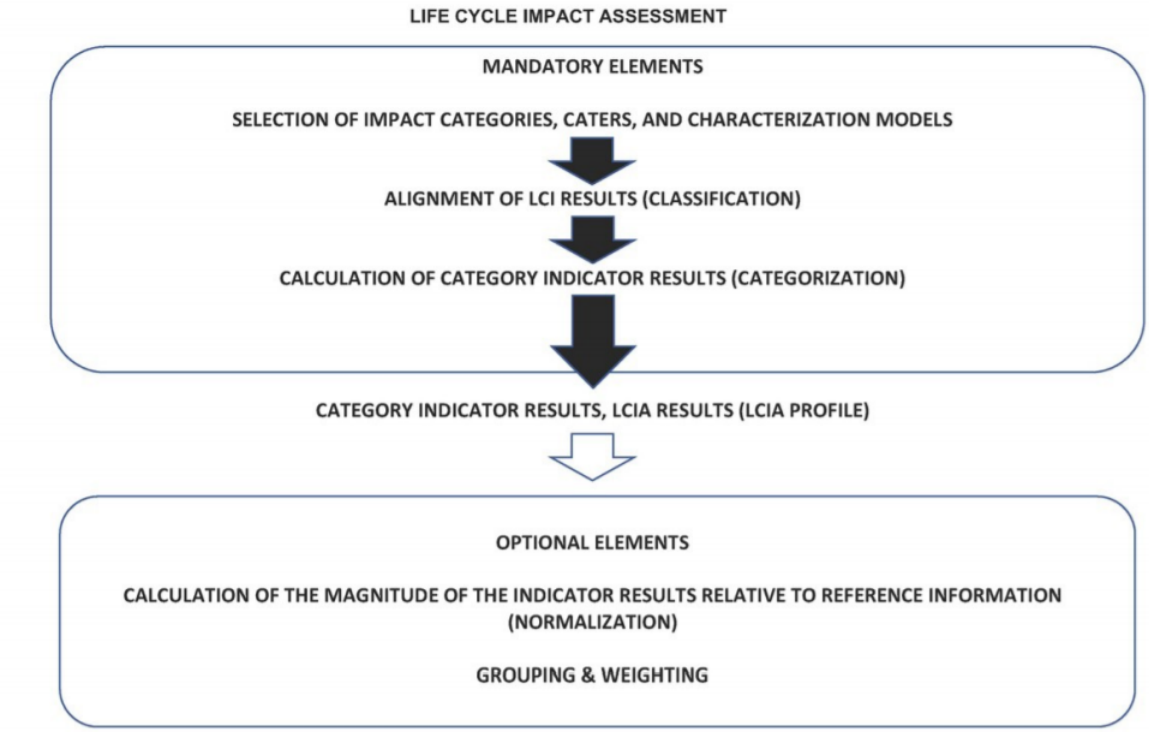


Figure 2.4.2.2.4: Impact assessment ISO mandatory and optional steps

This page titled 2.4.2.2: LCA/LCIA Concepts is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **2.2: LCA/LCIA Concepts** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.2.3: Conclusions and Review

There are many tools to assist life cycle assessments. There are software and data packages designed for performing LCAs. No matter the form of the software, the use of some sort of LCA software and data management system is needed in LCAs. The life cycle inventory step of an LCA often requires a large data set listing hundreds of emissions to the environment. Keeping track of these flows manually is arduous, so LCA software is designed to manage these flows and perform specific functions such as impact assessments based on the inventory as well as uncertainty analysis.

There is a large list of LCA software emerging with various features. A basic overview of how data and LCA software will first be provided then a list of software packages. LCA software can be split into several components:

- The software package
- Data sets
- Life cycle impact assessment (LCIA) methods

Software packages such as SimaPro, openLCA, and Gabi are frameworks or calculators that keep track of data and performs intensive numerical calculations. With the many flows and detailed 49 data, much effort has been invested in creating efficient calculation methods to speed up analysis time. This framework, however, is not useful without inventory data. There are many premade datasets provided from sources such as Ecoinvent, Gabi and United States Department of Agriculture (USDA) that contain previous life cycle inventory results for various chemicals, materials, energy, services, and waste treatment processes. LCA software can access this previously developed data and include a chemical or other process from a dataset in their LCA without needing to perform an entire LCA on that particular material or process. This fundamental aspect of LCA, the leveraging of previous study results for new studies, is a key benefit of LCA software and can save countless hours on the LCI step. LCIA methods are procedures and conversions that are used in performing a LCIA such as global warming potential characterizations and weighting methods. There are many accepted LCIA methods that calculate LCA results using different impact categories, types of impacts, and weighting methods.

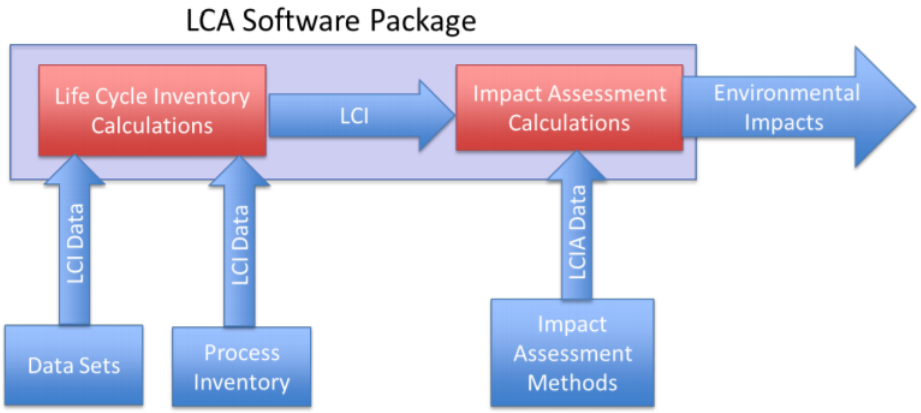


Figure 2.4.2.3.1: Life cycle assessment software structure

Figure 2.4.2.3.1 visually depicts how the different components of LCA software and data interact. The life cycle inventory step requires data from datasets (e.g., Ecoinvent) and primary data gathered by the LCA practitioner surrounding the process or product under analysis. The combination of these two types of data with the use of LCA software calculations gives an LCI. The LCI data can then be used to perform an impact assessment using the LCIA methods (e.g., TRACI).

Table 2.4.2.3.1: Three common LCA Software package options

Software	Licensing	Datasets	Software Features	Website
openLCA	Open source and free	Ecoinvent, Gabi, USLCI, CML and others	Fast calculation engine, easily share models, no yearly subscription, process based with transparent data, used for USDA digital commons LCA data development	www.openLCA.org
SimaPro	Paid licensing	Ecoinvent, USLCI, CML and others	Process based with transparent data, good customer support, robust uncertainty analysis	www.pre-sustainability.com/simapro
Gabi	Paid Licensing	Gabi dataset, Ecoinvent, USLCI	Robust dataset, visual process flow, based modeling, ease of use	www.gabi-software.com

Review Questions

1. Define the opportunities that could benefit from an LCA and why?
2. Is LCA going to be completely exhaustive in any of the various gate scenarios for determination of impacts?
3. What software is available for LCA?
4. Come up with a process and define 10 or less flows to allow you to calculate impacts.

Further Reading

1. Margni and Curran 2012 (See “Chapter 2 Supplementary Material in Moodle site)
2. Novel Screening Technique: Integrated Combinatorial Green Chemistry & Life Cycle Analysis (CGC-LCA): ojs.cnr.ncsu.edu/index.php/Bi...ning_Technique

---

This page titled 2.4.2.3: Conclusions and Review is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **2.3: Conclusions and Review** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

SECTION OVERVIEW

2.4.3: Hazards

2.4.3.1: Introduction to Hazards of Chemistry

2.4.3.2: Hazard Concepts

2.4.3.3: Case Study - Badger Army Ammunition Plant

2.4.3.4: Green Technologies for Safer Chemical Production

2.4.3.5: Conclusions and Review Questions

This page titled 2.4.3: Hazards is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

2.4.3.1: Introduction to Hazards of Chemistry

Hazards Introduction

What we as a population must realize is that no substance in and of itself is a poison or a remedy. It is the dosage that defines the activity of the substance. For example, let's look at the concept of "hormesis" as a precursor to the concept of poison/remedy by dosage.

We all have heard at one time or another the phrase, "What doesn't kill you makes you stronger". This phrase contains truth and contains at its essence the theory of hormesis: when organisms are exposed to low levels of stressors or toxins, they become more resistant to larger levels of stressors or toxins. This theory has been met with skepticism. Recently, however, biologists have put together a molecular explanation of its function and it has finally been accepted as a fundamental principle of biomedicine.

For example, exposing mice to low levels of gamma ray radiation before irradiating them with high levels actually decreases their likelihood of cancer. Similarly, when dioxin is given to rats we find the same situation. However, the biochemical mechanisms are not well understood. It is believed that a low dose of a toxin can trigger repair mechanisms that are efficient enough to not only neutralize the toxin, but repair other defects not caused by the toxin.

Thus, hormesis is nature's way of dealing with harmful agents; in fact, antibodies are a natural consequence of hormesis. However, the toxins/poisons that were once absolute, are NO LONGER. For example, thalidomide was found to be a very dangerous chemical for embryo development, but has recently found great promise in a number of ailments according to the Mayo Clinic including HIV, skin lesions, and multiple myeloma (please see: http://www.mayoclinic.org/diseasesco...e/art-20046534). Such a fact is outstanding considering the horrific aftermath (a few photographs shown below(Figure 2.4.3.1.1) of its use in the middle part of last century:



Figure 2.4.3.1.1: A child disfigured because of thalidomide https://www.flickr.com/photos/22719239@N04/2241322031



Figure 2.4.3.1.2: A few adults with disfigurements directly caused by thalidomide intake by the natural mother https://www.flickr.com/photos/bowbrick/3525018934

In fact, the list of former “pure” toxins is extremely interesting: snake venom, bacteria (botulin), fungi (penicillin), leeches (Hirudin), maggots (gangrene), etc. The toxic aspect notwithstanding, we all live in the wake of a world and society that is rife with potential hazards.

Types of hazards

A hazard is “threat” to life, health, property, or environment. These can come in many forms, but they are classified according to their modalities or nature of operation. The modalities of hazards are the following:

- Dormant: Has the potential, but nothing currently can be affected. This modality is typified by the “volcano” scenario – a volcano that is no longer showing any signs of activity or imminent threat but is lying “dormant”. There is no immediate and pressing issue based on human perception.
- Armed: Has the potential, and something can be affected. This modality is given by a person holding a gun in the midst of a war or other aggressive situation. The gun has the potential to affect life, limb, or other vital function, but it is not yet doing so, although the intention is there
- Active: Currently on-going event (something is being affected). Finally, this is the modality that is actually affecting life, limb, or some other vital function. A volcano that is spewing, a gun that is being fired, a fire that is consuming a building, radiation that is leaking, etc. These are the situations that are causing harm.

Within these modalities of hazards, hazards can be further refined as to their types. The following types of hazards classify the threats to life:

- Physical: Condition or situations that cause the body physical harm, e.g., a bullet that is entering into a human being.
- Chemical: Substances that cause harm or damage to the body, property or the environment, e.g., liquid oxygen converting to gaseous oxygen within a closed container (bomb).
- Biological: Biological agents that cause harm to the body, e.g., anthrax bacteria.
- Psychological: Stress affecting the mental state, e.g., the knowledge that the trajectory of a fivemile wide asteroid coincides with earth.
- Radiation: Electromagnetic radiation that harms or damages biological organisms.

This page titled 2.4.3.1: Introduction to Hazards of Chemistry is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1: Introduction to Hazards of Chemistry** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.3.2: Hazard Concepts

Concepts

Globally Harmonized System

To best classify hazards, the concept of GHS (Globally Harmonized System) has been introduced. Figure 2.4.3.2.1 demonstrates the categories of hazards:



Figure 2.4.3.2.1: The globally harmonized system (GHS) that attempts to specifically categorize the general types of threats in society. Please see emergency.cdc.gov/agent/agent...emcategory.asp for more information; commons.wikimedia.org/wiki/File:GHS_HAZCOM_Safety_Labels.jpg

Notice that each logo attempts to demonstrate pictorially the type of hazard represented by a specific chemical or material. For example, notice Figure 2.4.3.2.2 below. This particular chemical's hazard level could be listed as "gases under pressure" and "explosives". The intensity of its peculiar hazards is evaluated according to the pressure and the explosive nature of the gas.



Figure 2.4.3.2.2: A tank of propane gas commons.wikimedia.org/wiki/F...oamFatale_Pres sure_Vessel.jpg

Exposure and Categories of Hazardous Substances

Our overall sensitivity to hazards, again, depends on a number of factors that not only include the typical factors of levels, LDs, type of toxin, modality, etc., but also to our idiosyncratic immunity or responsiveness.

We have been accustomed to a number of wonderful creature comforts in life that have only come about because of the power, versatility, and creativity inherent in the chemical enterprise. For example, phosgene, a notorious chemical warfare agent in the Great War, is now used as a precursor for the manufacture of a number of items including polyurethane. Shown below in Figure 2.4.3.2.3 is a representation of phosgene:

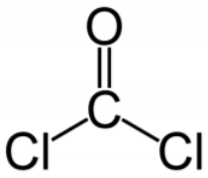


Figure 2.4.3.2.3: A chemical representation of phosgene. http://www.toxipedia.org/display/toxipedia/Phosgene

Lab Safety

Lab safety is an essential part of working with hazardous materials. It is important to follow good lab practice techniques and to know how to deal with situations should they occur. In the following video by Chemistry crash courses, some good lab techniques and safety guidelines are discussed. Lab Techniques and safety by crash course: https://www.youtube.com/watch?v=VRWRmIEHr3A

Reference: Green, H. Crashcourse. (2013, July 8). Lab Techniques & Safetey: Crash Course Chemistry#21. Retrieved from https://www.youtube.com/watch?v=VRWRmIEHr3A

Lethal Dose

Lethal dose (LD50) is the amount of any ingested or interfering substance that kills 50% of a test 58 sample. It is expressed in mg/kg, or milligrams of substance per kilogram of body weight. In toxicology, it is also referred to as the median lethal dose that refers specifically to a toxin, radiation, or pathogen. The lower the LD50, the more toxic is the item being measured for toxicity. It can be considered a pragmatic approach to toxicity exposure levels because in general, toxicity does NOT always scale with body mass. The choice of 50% lethality as the gold standard avoids ambiguity because it measures in the extremes and reduces the amount of testing. However, such a fact also means that LD50 is not the lethal dose for all subjects; in other words, some may be killed by much less. Measures such as “LD1” and “LD99” (dosages required to kill 1% or 99%, respectively, of the test population)

are sometimes used. Shown below is a chart of sample LD50 taken from Wikipedia (en.Wikipedia.org/wiki/Median_lethal_dose) with active links to allow further investigation of the substances whose measurements are given.

Substance	LD_{50} (LC_{50})	LD_{50}: g/kg (LC_{50} g/L) standardized
Water	>90 g/kg	>90
Pentaborane	<50 mg/kg	<0.05
Cobalt(II)chloride	80 mg/kg	0.08
Metallic Arsenic	763 mg/kg	0.763
Cadmium oxide	72 mg/kg	0.072
Sucrose (table sugar)	29,700 mg/kg	29.7
Monosodium glutamate (MSG)	16,600 mg/kg	16.6
Vitamin C (ascorbic acid)	11,900 mg/kg	11.9
Urea	8,471 mg/kg	8.471
Cyanuric acid	7,700 mg/kg	7.7
Cadmium sulfide	7,080 mg/kg	7.08
ethanol (grain alcohol)	7,060 mg/kg	7.06
sodium isopropyl methylphosphonic acid (IMPA, metabolite of sarin)	6,860 mg/kg	6.86
Melamine	6,000 mg/kg	6.00
Melamine cyanurate	4,100 mg/kg	4.1
Venom of Brazilian Wandering Spider	134 µg/kg	0.000134
Venom of Inland Taipan (Australian snake)	25 µg/kg	0.000025
Ricin	22 µg/kg 22-30 mg/kg	0.000022 0.02
2,3,7,8- Tetrachlorodibenzodioxin (TCD-D, a dioxin)	20 µg/kg	0.000002

Sources for toxicity information:

- http://www.atsdr.cdc.gov/
- http://cfpub.epa.gov/ecotox/
- toxnet.nlm.nih.gov/

Figure 2.4.3.2.4: A Brazilian Wandering Spider commons.wikimedia.org/w/index.php?curid=332527

Notice that the venom of the Brazilian Wandering Spider is particularly potent. A ten thousandth of a g/kg or ~5mg would kill a normal sized female. Luckily the venom discharged per bite is quite small. Nevertheless, the point is that the vector (or venom) represents a toxin that is of sufficient lethality that it can cause great harm or injury to a human being.

This page titled 2.4.3.2: Hazard Concepts is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **3.2: Hazard Concepts** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.3.3: Case Study - Badger Army Ammunition Plant

Badger Army Ammunition Plant (BAAP) Case Study

Case Study for Hazardous Waste



Figure 2.4.3.3.1: A scene from the BAAP waste cleanup site https://dma.wi.gov/DMA/news/2015news/15042

Environmental cleanup of the 7,400-acre Badger Army Ammunition Plant will be greater than $250 million. This is one of the 40 contaminated military sites in Wisconsin which the Defense Environmental Restoration Account cites as the most contaminated. In fact, 32 areas are polluted with dangerous levels of solvents, metals and explosive/incendiary waste. The water beneath the plant is contaminated with mutagenic chemicals that include carbon tetrachloride, trichloroethylene and dinitrotoluenes. One part of the site, known as the Propellant Burning Grounds, is the source of a three-mile plume of contaminated groundwater that has migrated offsite, completely contaminating private drinking water wells in addition to the Wisconsin River.

Elevated Cancer Rates

In 1990, the Wisconsin Division of Health conducted a health survey. It concluded that communities near the Badger plant have a significantly higher incidence of cancer and deaths. In spite of these alarming findings, the State refused to take any action. In 1995, the Division of Health responded because of pressure from CSWAB (Citizens for Safe Water Around Badger) and reopened the community health study.

On October 26, 1998, CSWAB concluded “there was entirely inadequate contact with our community – the population being studied.” Prior to September 1998, no press releases were published, no public meetings were held, and no interviews were conducted. Despite several requests, virtually no resources were devoted to interviewing residents about current health problems and concerns regarding their exposure to air, dust, emissions, and surface soils. Assessment of risk from cleanup activities was also absent. CSWAB also determined that the Wisconsin DOH focused on death studies when many health problems and community concerns have been nonlethal, such as respiratory illnesses or reproductive problems.

CSWAB appealed to ATSDR (Agency for Toxic Substances and Disease Registry in D.C.) to discuss the lack of adequate community participation in this and similar health assessments across the nation. The Wisconsin Department of Health conducted the assessment under a cooperative agreement with ATSDR.

Cleanup Plans Abandoned

The U.S. Army proposed to abandon and severely reduce cleanup of two priority areas within the plant. The Settling Ponds and Spoils Disposal area — a series of lagoons that run the length of the 7,000-acre facility — is contaminated with high levels of lead and dinitrotoluenes.

Maintenance of the Badger plant costs in excess of $17 million per year. In 1991, only $3 million was allocated for environmental studies. Since 1975, there have been over 56 chemical spills and incidents. Moreover, there is no national strategic need for maintaining the Badger plant. A February 20, 1997 Government Accounting Office report concludes BAAP and three other military plants could be eliminated “because alternative sources exist…to provide the capabilities these plants provide.”

This page titled 2.4.3.3: Case Study - Badger Army Ammunition Plant is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **3.3: Case Study - Badger Army Ammunition Plant** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.3.4: Green Technologies for Safer Chemical Production

Since the fallout from the Bhopal tragedy, efforts have centered on avoiding storage of methyl isocyanate (MIC). The final product, Sevin, is no longer being manufactured by a two-step process. The process now consists of no longer reacting naphthol with MIC (one step), but sequentially with phosgene and methyl amine. The basis for the Bhopal tragedy was the reaction shown in Figure 2.4.3.4.1.

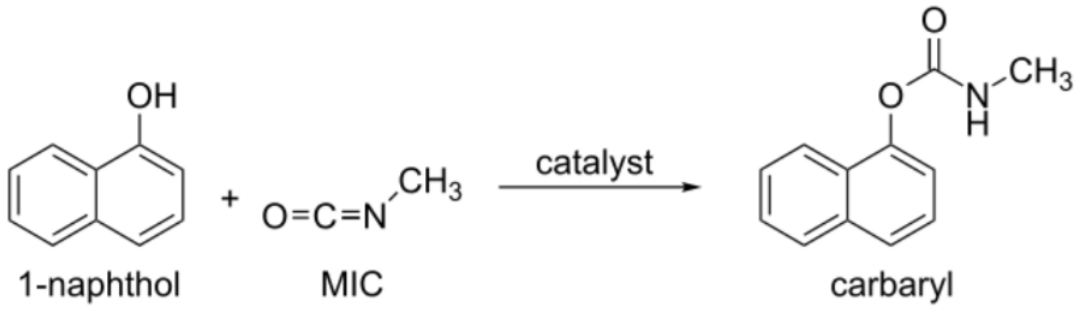


Figure 2.4.3.4.1: The original process that DuPont had in place to make Sevin (the commercial pesticide) through methyl isocyanate (MIC). http://www.wikiwand.com/en/Methyl_isocyanate

Polyurethanes

Polyurethane is a polymer composed of carbamate (urethane) linkages. Polyurethanes are traditionally formed by reacting a di- or polyisocyanate with a polyol (alcoholic polymer such as PEG, polyethylene glycol). Both the isocyanates and polyols used to make polyurethanes contain, on average, two or more functional groups (either on termini or within the molecule, hence being telechelic, i.e., a di-end-functional polymer where both ends possess the same functionality) per molecule. Recent efforts have been dedicated to minimizing the use of isocyanates to synthesize polyurethanes because isocyanates are toxic. Non-isocyanate-based polyurethanes (NIPUs), especially made from soybean oils, have recently been targeted as a new greener class of polyurethanes. Shown in Figure 2.4.3.4.2 is a molecular representation of polyurethane linkages (highlighted in blue). The carbamate linkages are composed of a central carbonyl moiety that has two heteroatoms attached to it – a nitrogen and an oxygen. From an organic perspective, it is an ester/amide hybridized molecular system. Interestingly, telechelic monomers such as adipic acid chloride (left telechelic monomer) and hexamethylene diamine (right telechelic monomer) such as shown in Figure 2.4.3.4.3.

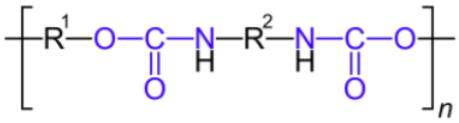


Figure 2.4.3.4.2: A simplified representation of urethane (carbamate) linkages in polyurethane polymers. de.Wikipedia.org/wiki/Polyurethane#/media/File:Polyurethane-allg.svg

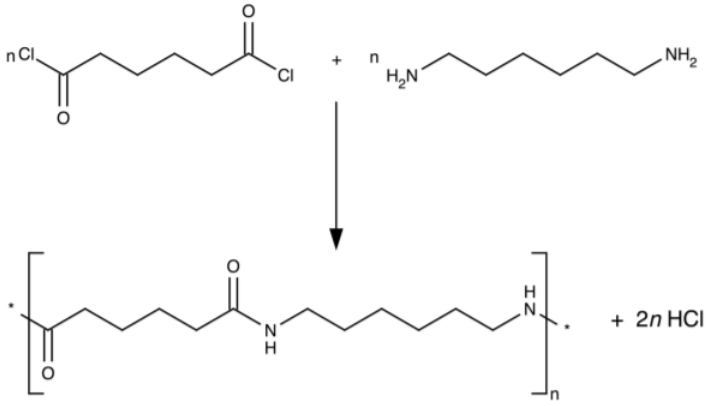


Figure 2.4.3.4.3: A simplified representation of the telechelic nature of the starting materials (monomers) that upon coupling form a telechelic polymer. commons.wikimedia.org/wiki/File:Nylon_synthesis_from_chloride_and_amine.svg

What is immediately noticeable is that the final polymer is a blend of two distinct monomers; thus, the final properties can be tailored by judicious (and discrete) choice of the monomers. More specifically, prepolymers (> 10-12 monomer units) of each monomer may be coupled to provide distinct segments having specific properties. For example, polyethylene glycol (PEG) is a polymer that is hydrophilic, soft, rubbery, and flows well. However, the phenyl-based di-isocyanate segment is much more rigid, tough, and non-stretchable. Therefore, the overall final physical and thermal properties of the polyurethane can be tuned. The opportunity to enhance the polymerization reactivity inherent for polyurethanes can be catalyzed by a non-nucleophilic base (such as DABCO – diazabicyclooctane) shown in Figure 2.4.3.4.4

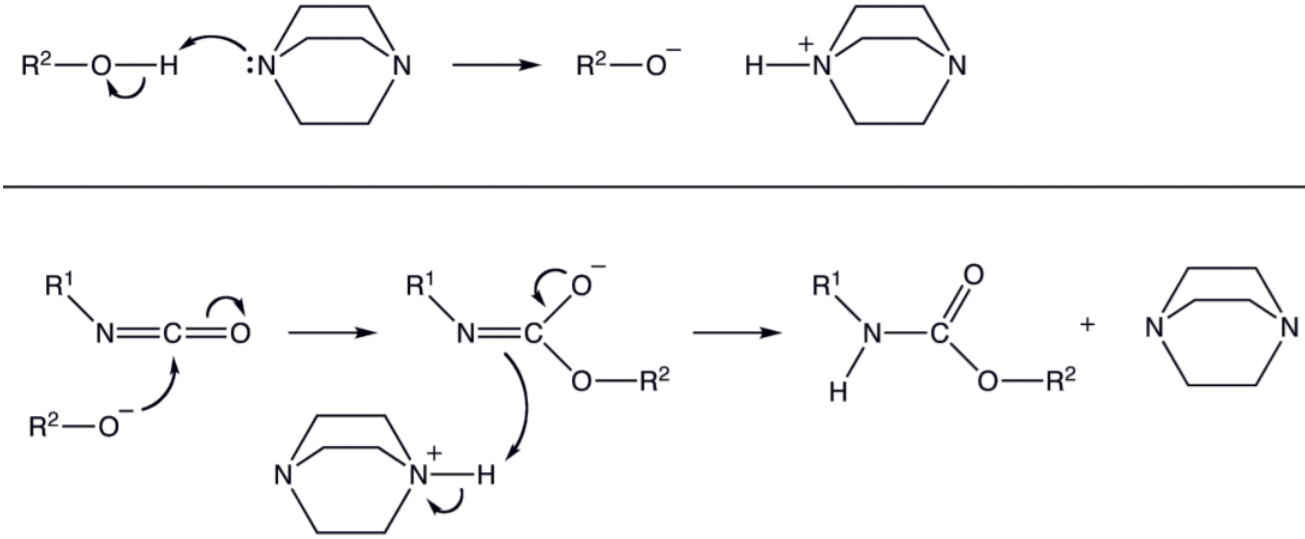


Figure 2.4.3.4.4: The molecular steps in the formation of a urethane bond by base catalysis.
http://www.wikiwand.com/en/Polyurethane

In Figure 2.4.3.4.4, DABCO is able to abstract a proton from an alcohol (say ethylene glycol) to allow for the nucleophilic reactivity at the cumulated carbon of the isocyanate.

Curtius Rearrangement to Form Isocyanates

An additional “greener” approach to forming isocyanates, an important class of starting material in the formation of many highly important materials, is through the Curtius Rearrangement (RAR) Reaction, an example of which is shown in Figure 2.4.3.4.5.

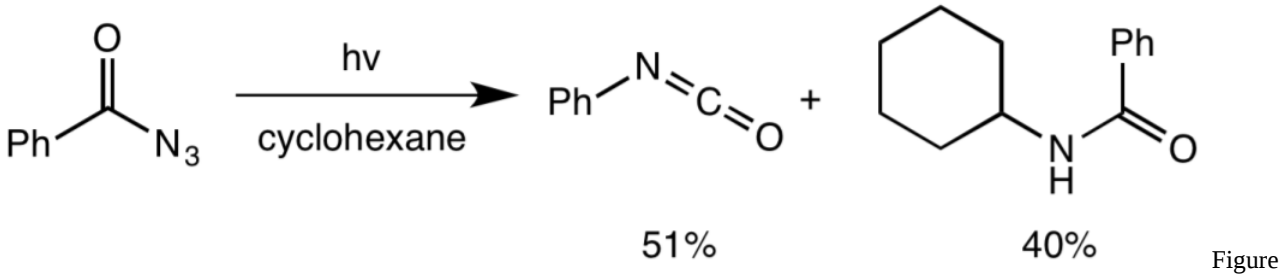
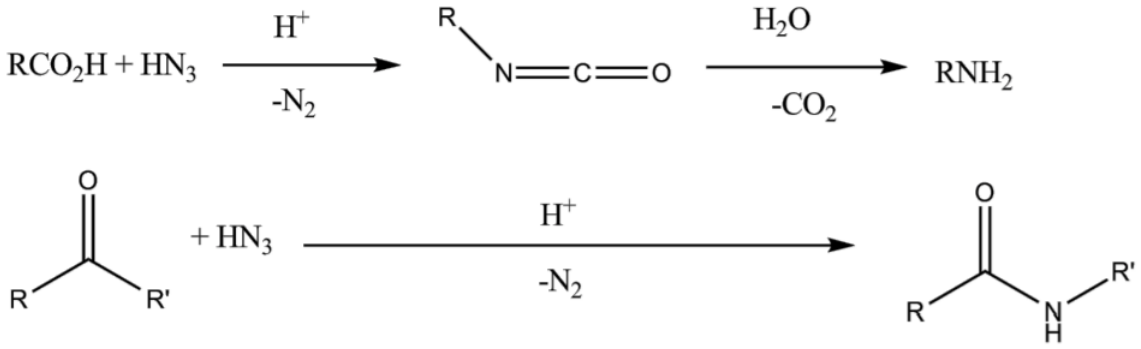


Figure 2.4.3.4.5: The Curtius RAR reaction http://www.wikiwand.com/en/Curtius_rearrangement

The Curtius Rearrangement is a thermal or photochemical decomposition starting from carboxylic azides (left structure in reaction above) to an isocyanate (first product in reaction above). In the above reaction, the solvent also plays the role of a reactant as shown by N atom insertion into cyclohexane via a radicaloid mechanism. These intermediates may be isolated, or their reaction or hydrolysis products can be obtained.

The reaction sequence that includes the subsequent reaction with water that leads to amines is called the Curtius Reaction. This reaction is similar to the Schmidt Reaction, shown below in Figure 2.4.3.4.6 (https://www.wikiwand.com/en/Schmidt_reaction), with acids that differs in that the acyl azide is prepared from the acyl halide and an azide salt.

Schmidt Reaction:



Mechanism of the Curtius Rearrangement:

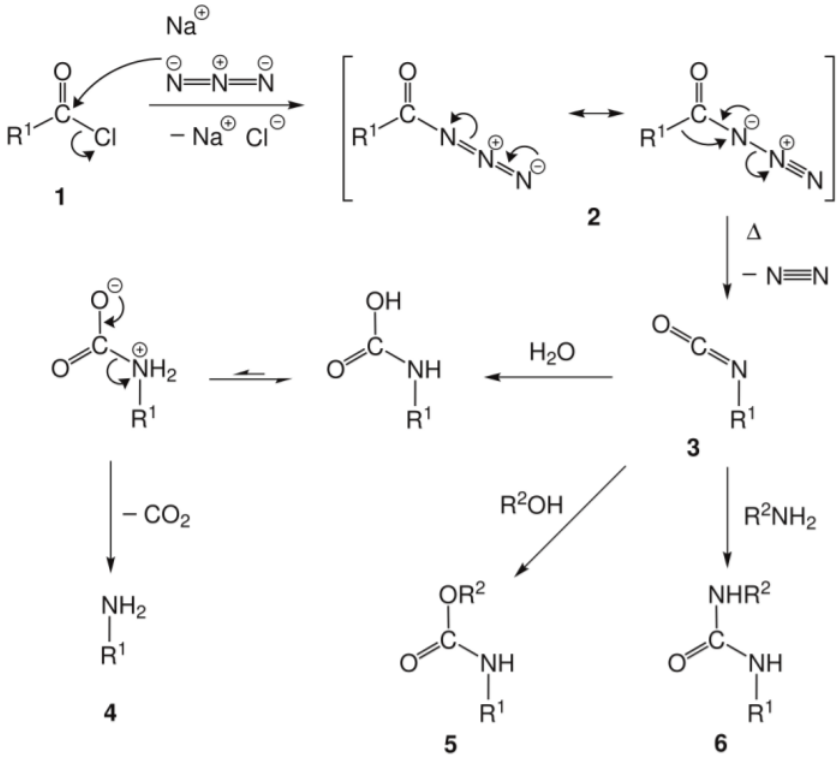


Figure 2.4.3.4.6: *Two different reactions to form amines, one showcasing the green chemistry principles*

For the Curtius RAR, the following steps take place: Starting Reagent 1 undergoes an electrophilic attack by an anionic azide molecule to produce an azide whose mesomerism is shown in brackets 2. It then undergoes decomposition under the appropriate stressor (heat, pressure, etc.), which leads to the isocyanate 3 with the loss of nitrogen (denitrogenation). This can react with water to yield a carbamate-like molecule (carbamic acid, urethane-like) which can spontaneously decompose thru decarboxylation to a primary amine 4. In the presence of an alcohol or an amine, 3 can yield an ester, 5, and an amide, 6, respectively. http://www.wikiwand.com/en/Curtius_rearrangement

---

This page titled 2.4.3.4: Green Technologies for Safer Chemical Production is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **3.4: Green Technologies for Safer Chemical Production** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.3.5: Conclusions and Review Questions

Conclusions

The opportunity to apply green chemistry to reducing the toxicity in industrial processes is of paramount importance for a sustainable future. In this chapter, we investigated the concept of hazards from the perspective of their meaning, nature, implications, control, and ultimate reengineering. Green chemical principles hold the key to ensuring that society provides a viable channel for all of our physical needs and maintains the environment.

Review Questions

1. Think about the additive impact of hazards from a LD50 perspective. Why would materials with different values not be additive in their cumulative effects on lethality? Can you think of two chemicals, however, whose lethality may be more than the sum of their individual values?
2. Would phosgene be more or less reactive as a function of relative humidity? Why or why not? In a wartime situation, for example, would an opposing army to whom phosgene is being directed favor a low humidity?
3. Can phosgene be reacted with sodium azide in a 1:2 molar ratio? If so, what would be the reaction and product? If not, why not?
4. What do these GHS labels tell us about ammonia? (See Figure 2.4.3.5.1)



Figure 2.4.3.5.1: A label for Ammonia https://www.general-data.com/about/blog/ghs-compliant-labels-what-are-essential-components

Further Reading

- blogs.ei.columbia.edu/2014/12...try-theanswer/
- https://play.google.com/books/reader...ec=frontcover& output=reader&hl=en_US&pg=GBS.PP1(Bretherick'sHandbookofReactive Chemical Hazards: Edition 4, L. Bretherick, October 22, 2013, Elsevier)
- http://ccc.chem.pitt.edu/wipf/Web/HCH.pdf (Hazardous Chemicals Handbook)

This page titled 2.4.3.5: Conclusions and Review Questions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **3.5: Conclusions and Review Questions** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

SECTION OVERVIEW

2.4.4: Alternative Solvents

In a very general sense, solvents are a class of chemical compounds that allow chemistry to occur. The concept of a solvent has significant ramifications because they serve as the matrix, medium, or carrier for solutes. They are necessary in a number of processes, reactions, and systems. We tend to think of a substance like “water” as a universal solvent because it is so useful in so many disciplines. Water cleans up everything, allows biochemical reactions to occur, is used in paints, coatings, and films, allows cooking to occur (or else everything would catch fire), and provides lubrication and ease of movement for a great many devices.

2.4.4.1: Introduction to Solvents

2.4.4.2: Solvent Concepts

2.4.4.3: Solvent Alternatives

2.4.4.4: Conclusions and Review

This page titled 2.4.4: Alternative Solvents is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

2.4.4.1: Introduction to Solvents

In a very general sense, solvents are a class of chemical compounds that allow chemistry to occur. The concept of a solvent has significant ramifications because they serve as the matrix, medium, or carrier for solutes. They are necessary in a number of processes, reactions, and systems. We tend to think of a substance like “water” as a universal solvent because it is so useful in so many disciplines. Water cleans up everything, allows biochemical reactions to occur, is used in paints, coatings, and films, allows cooking to occur (or else everything would catch fire), and provides lubrication and ease of movement for a great many devices.

Interestingly, in the business world, the word “solvent” has another meaning: A state of financial soundness characterized by the ability of an entity to meet its monetary obligations when they fall due. This latter definition does indeed apply in a certain sense to the solvent concept we are espousing in this chapter. A solvent ($olvent) carries materials for a reaction or specific function in the same sense that solvent from a financial sense carries funding/money to meet its obligations. Thus, using an analogy:

Solvent : Solute :: $olvent : Money

Chemically, a solvent will dissolve or “solvate” a solute. What does that mean? It means that solvent molecules will surround the solute in such a way that a solution is formed; in other words, a homogeneous system is generated in which the solute is part of/indistinguishable from the solvent network. A solute/solution conceptcan be visualized in the Figure below:

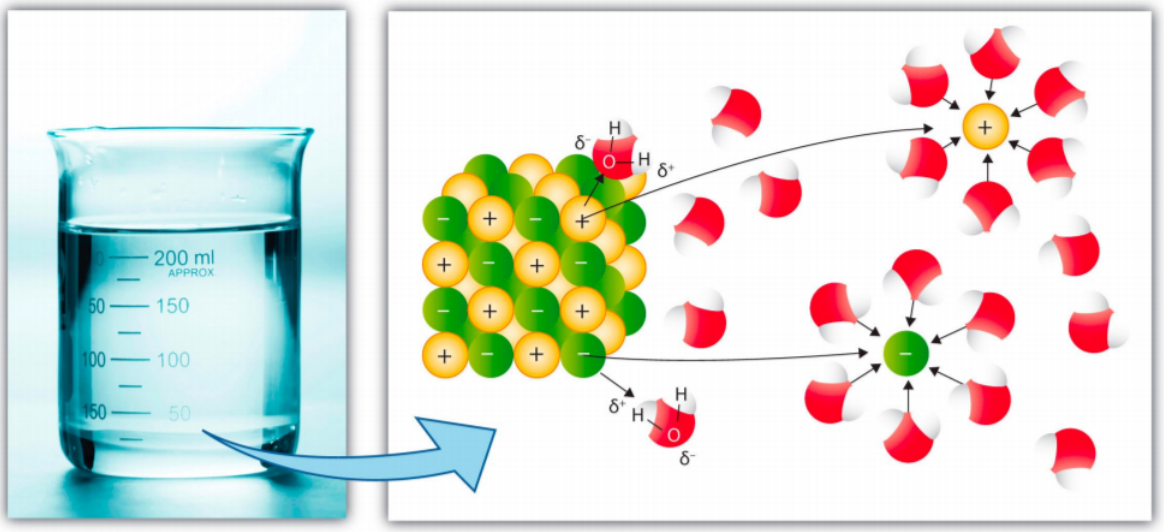


Figure 2.4.4.1.1: Shown is a simplified representatino of how a solvent such as water can "encapsulate" or cause a solute to go into solution (or be washed away). https://commons.wikimedia.org/wiki/File:Sodium_chloride_dissolution.jpg

Solubility

The ability of one a compound to dissolve in some other (likely different) compound is termed “solubility”. Miscibility is another term that characterizes the facility of compound A to dissolve in compound B. When the two compounds can completely dissolve or combine to form a homogeneous solution, the two liquids are said to be miscible. Two that can never blend well enough to form a solution are called immiscible.

All solutions have a positive entropy of mixing whereas the interactions between different compounds may or may not be energetically favored. If interactions are unfavorable, then the free energy decreases with increasing solute concentration. The energy loss may at some point in time outweighs the entropy gain, and no more solute can be dissolved – this is a condition solution in which the solution is said to be saturated. This condition can change with different environmental factors, such temperature, pressure, and purity of the system. A supersaturated solution can be prepared by raising the solubility (e.g., increasing temperature) to dissolve more solute, and then lowering it by cooling. However, most gases and several compounds exhibit solubilities that tend to drop with raised temperatures. The solubility of liquids in liquids, however, is less temperaturesensitive than solids or gases.

This page titled 2.4.4.1: Introduction to Solvents is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **4.1: Introduction to Solvents** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.4.2: Solvent Concepts

According to GC Principle No. 5, “The use of auxiliary substances (solvents, separation agents, etc.) should be made unnecessary wherever possible, and innocuous when used.” This statement signifies or upholds the dogma in GC that simplification to execute a specific transformation should be observed. In other words, if we can find a way to not use a solvent, let’s not use it! Although in our modern society such a stance is almost untenable, the ideal nevertheless is what we as a society should contend to achieve. For example, as shown in Figure 2.4.4.2.1, we have an overwhelming influx of medicines, drugs, etc., within our society to promote our health. We typically take these concoctions with water, in emulsions, in solutions, in suspensions, etc.



Figure 2.4.4.2.1: The gamut of medicines and its varieties in our culture and life is staggering. pixabay.com/en/health-medicine-tablets-846862/

Of course, taking a medicine without solution is a difficult proposition, but there are a number of ways to do it:

- Dry swallowing or sublingually (e.g., nitroglycerin tablets for angina);
- Patch application (microneedles or high concentration, usually done dermally);
- Inhaled by a mist/spray(e.g., gaseous phase);
- Via a tube (e.g., stomach tube or intravenous).

Nevertheless, the use of solvents and solvent-based systems are deeply ingrained in our culture and society. The reasons are many, but they tend to be tradition, ease of use, reduced cost, and convenience. Doing without solvents tends to require much more creativity and overall planning.

Reaction Energy Coordinate

Typically, a reaction profile or surface (where and how a reaction proceeds according to an energy perspective) provides sufficient information to understand the pathways (independent of final energy states) necessary to achieve a forward reaction. For example, the coordinate shown below illustrates the SN2 reaction between t-butyl chloride and hydroxide. Note that most of the variables shown relate to the energy constraints or parameters associated with the reaction. The reaction has an inherent activation barrier for the forward direction. This barrier characterized by an energy differential input (DE_1^1), relates to the difficulty to accessing a transition state (^{1}) that must be accessed for the forward reaction.

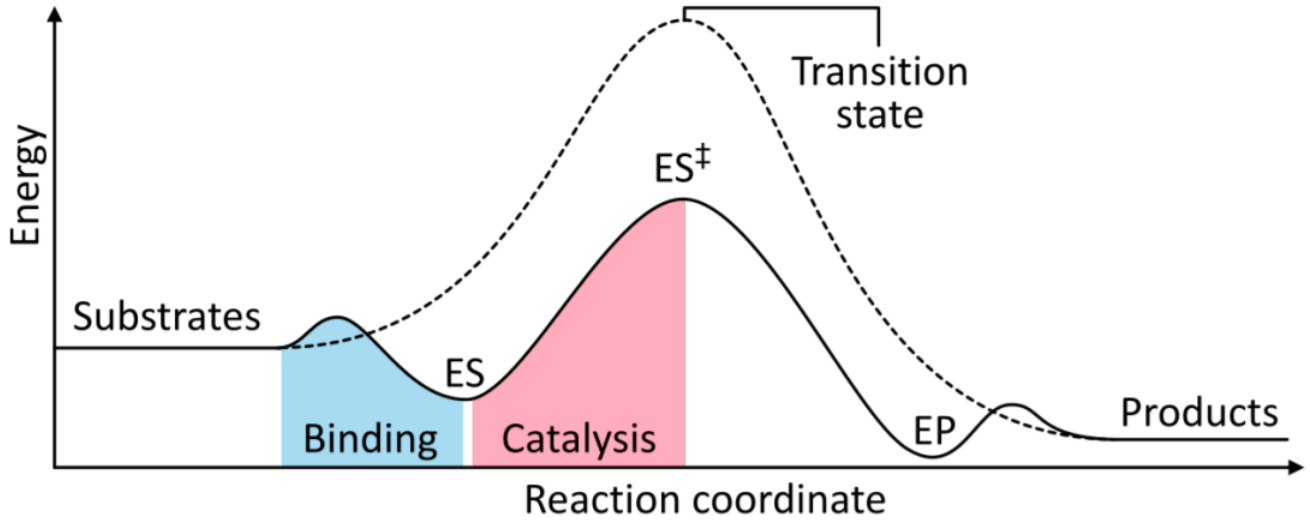


Figure 2.4.4.2.2 *A typical representation of the nergy surfaces obtained during the transformation of an organic halide in a classic S_N2 (bimolecular nucleophilic substitution) reaction.*
https://commons.wikimedia.org/wiki/File:Enzyme_catalysis_energy_levels_2.svg

Such an energy input, however, can be accommodated or lessened by a number of factors, viz., the use of a catalyst which in general has shown the ability to reduce the energetic considerations to access that state by virtue of its capacity to engage in structural or energetic interactions with the starting material and strongly encourage the formation of the transition state. The downhill nature of the coordinate surface is akin to rolling a ball whose potential energy conversion to kinetic energy is extremely facile. Thus, an intermediate is achieved (the trialkyl carbocation in Figure 2.4.4.2.2) that has a measurable lifetime and properties. Although it is by its fundamental character unstable, it nevertheless exists to then engage in further reactions. Notice that it still at an energy surface that is much higher to the starting energy state. Such a phenomenon is then predictive of its future course. In fact, as you move closer to the product along the abscissa of the reaction coordinate surface, the intermediate resembles in a number of properties the nature of the final product. In this case, the intermediate has to a great extent a number of the features of both the starting materials and the product. Indeed, the next activation barrier that must be crossed to access the final product is much lower in height than the normal activation barrier indicating the gradual easing of the transition of starting material(s) to product(s).

Solvent Categories

There are a number of solvents that from a green chemistry perspective must be dealt with judiciously. The following are representative solvent classes:

- Hydrocarbons
- Halogenated hydrocarbons
- Aromatic hydrocarbons
- Alcohols
- Ethers
- Aprotic solvents

Each of the above have their pros & cons in terms of environmental benefit/issue, economics, and social justice. Each may do different things; for example, the aprotic solvent class is one where there is no chance for the solvent to provide a proton to a reaction it is hosting. Alcohols, on the other hand, can easily do that and should not be used in reactions where water/protons could quench a reaction or trigger a violent reaction.

Methyl Soyate

Methyl soyate is a biobased solvent that is a mixture of long-chain fatty acid methyl esters. More information on its specific properties and potential uses can be found in the following tract:

https://www.yumpu.com/en/document/view/10362159/the-formulary-guide-for-methyl-soyate-soy-new-uses

In general, it can be used to clean countertops, pretreat fabric stains, clean concrete, degreaser, graffiti remover (with ethyl lactate and surfactants), paint stripper, mastic remover, varnish remover, deinker, asphalt remover, and waterless hand cleaner. In addition to the soyate, there are a number of other emerging biobased cleaners/solvents such as shown below:

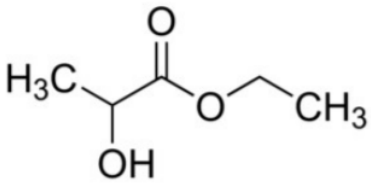


Figure 2.4.4.2.3: A representation of ethyl lactate, also known as lactic acid ester, which is a monobasic ester formed from lactic acid and ethanol, that can be used as a solvent. It is considered biodegradable and can be used as a water-rinseable degreaser. https://www.wikiwand.com/en/Ethyl_lactate

Please see http://pubs.rsc.org/en/Content/ArticleLanding/2011/GC/c1gc15523g#!divAbstract for more information on this marvelous solvent.

This page titled 2.4.4.2: Solvent Concepts is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **4.2: Solvent Concepts** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.4.3: Solvent Alternatives

Alternative Solvent Systems or Modes

Gas Phase

The gas phase is a very useful modality to allow reactions to occur because (like the hydrophobic effect) it forces reactions or processes to occur through non-solvent-mediated channels. For example, the production of methanol can be done in the gas phase by reaction of “syn” gas (hydrogen and carbon monoxide) with ZnO as the solid catalyst for the reaction to occur.

No Solvents

It is altogether possible and highly desirable to use the starting materials themselves for the reaction of interest. This is doable if one of the reactants is a liquid that can allow the other(s) to dissolve into it. This has been shown by taking a p-xylene and reacting it with oxygen to make the terephthalic acid:

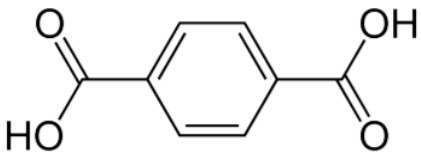


Figure 2.4.4.3.1 *A molecular representation of terephthalic acid* https://www.wikiwand.com/en/Terephthalic_acid

The acid in Figure 2.4.4.3.1 can then be used directly with ethylene glycol to synthesize polyethylene terephthalate:

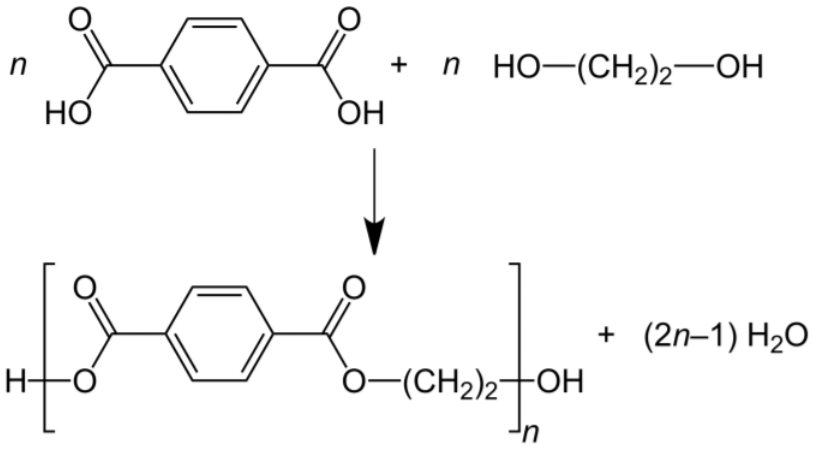


Figure 2.4.4.3.2 *The reaction of terephthalic acid with ethylene glycol leading to the production of the polymer.* https://www.wikiwand.com/en/Polyethylene_terephthalate

Melt State

It is possible to combine two solids to provide a composition that can achieve a eutectic point, or melts/solidifies at a single temperature lower than the melting points of the separate constituents or of any other mixture of them. At the eutectic point, you can achieve an isotropic blend of dissolved materials such as evidence in Figure 2.4.4.3.3.

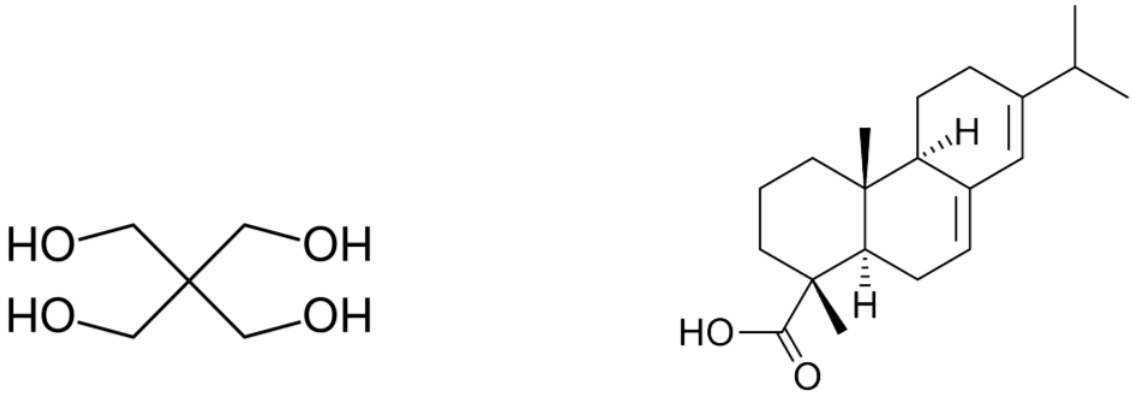


Figure 2.4.4.3.3: A mixture of pentaerythritol (left) and abietic acid (right) can blend in a melt (eutectic) that can provide a pentaester product. https://www.wikiwand.com/pt/Pentaeritritol; https://www.wikiwand.com/en/Abietic_acid

Triggered Solid State Reactions

It is possible to induce a chemical reaction outside of the melt by introducing a trigger such as acid, ultrasound, grinding, light, etc. Such triggers take the place of solvent-mediated reaction stabilization. For example, the following reactions are amenable to such triggers:

Michael Reaction:

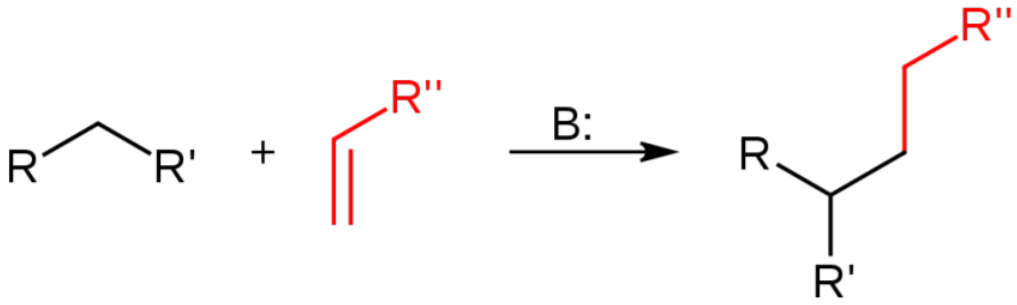


Figure 2.4.4.3.4: A simplified representation of the Micheal reaction. The conjugation of alkanes to unsaturated substrates is a key step for the synthesis of various complex natural products, antibiotics, and chiral compounds. http://www.wikiwand.com/en/Michael_reaction

In addition, it is possible to do the reaction using a catalyst such as alumina with microwave induction as shown in Figure 4-9:

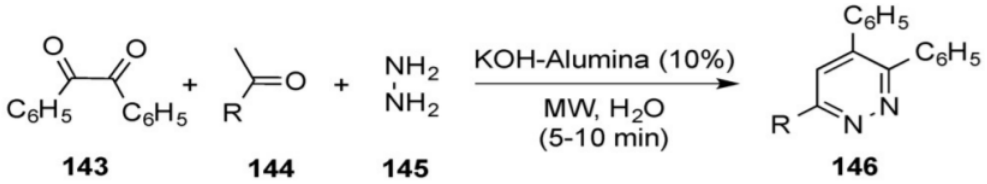


Figure 4-9 proceeded very efficiently using potassium hydroxide on alumina (KOH-alumina) as a mild, efficient, and recyclable catalyst. This approach gave pyridazines with high yields (73%–89%) after only several minutes.

There are a number of other reactions that are similar in their reactivity as a function of triggers in water (non-organic) and without catalysts.

Baeyer-Villiger:

For example, the Baeyer-Villiger Reaction is an important reaction in making an ester or lactone

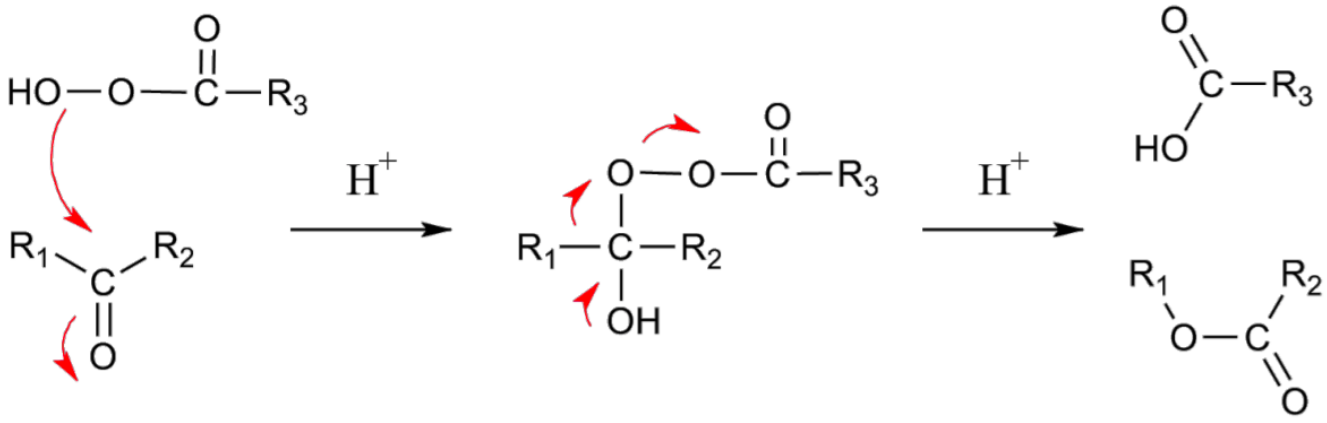


Figure 2.4.4.3.6: A depiction of the Baeyer-Villiger reactions with electron pushing arrows. commons.wikimedia.org/wiki/File:Baeyer- Villiger_oxidation_(mechanism).png

The Baeyer-Villiger Oxidation is the oxidative cleavage of a carbon-carbon bond adjacent to a carbonyl to converts ketones to esters and cyclic ketones to lactones. It may vw carried out with peracids, such as m-CBPA, or with hydrogen peroxide and a Lewis acid.

Benzilic Acid Rearrangement

1,2-Diketones undergo a rearrangement in the presence of a strong base to yield αhydroxycarboxylic acids. The best yields are obtained when the diketones do not have enolizable protons. Shown below in Figure 2.4.4.3.7 is a representation of the mechanism of the reaction.

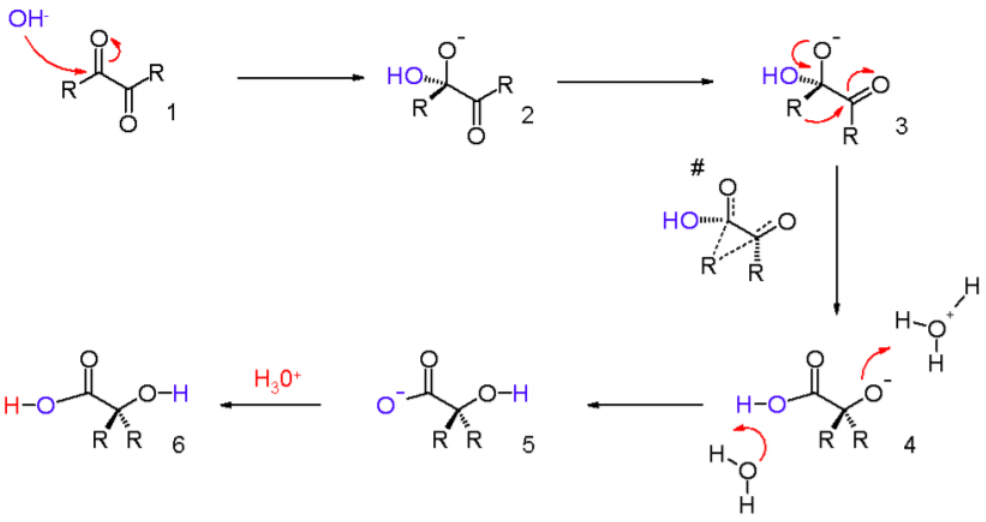


Figure 2.4.4.3.7: The mechanism of the Benzillic Acid Rearrangement. https://commons.wikimedia.org/wiki/File:Benzilicacidrearrangement.png

Benzilic acid rearrangement has traditionally been conducted by heating benzyl derivatives and alkali metal hydroxides (KOH) in aqueous organic solvent. However, the rearrangements proceed more efficiently and quickly in the solid state.

Williamson Ether Synthesis

The Williamson Ether Synthesis is an organic reaction forming an ether from an organohalide and a deprotonated alcohol (alkoxide) that is done in organic solvents. However, a very rapid synthesis of symmetrical and asymmetrical ethers in “dry” media under microwave radiation has been reported (http://www.cyfronet.krakow.pl/~pcbogdal/alcohol/). The reaction was carried out by mixing an alcohol with 50% excess of an alkyl halide and catalytic amount of tetrabutylammonium bromide, adsorbed onto potassium carbonate or a mixture of potassium carbonate and potassium hydroxide, and irradiated in open conditions in a domestic microwave oven for 45-100 s. In the absence of the ammonium salt, ethers were not detected or were very low yield.

Friedel-Crafts Reaction

The opportunity to perform acylation or alkylation of an aromatic nucleus is a very important transformation in organic chemistry. The principal tool we have for such a transformation is the Friedel-Crafts, a classical reaction dating back to the 19th century, can be done in a green way to avoid the need for acid chlorides, Lewis acids, and hydrochloric acid waste products. Shown below in Figure 2.4.4.3.8 is a representation of the photo-Friedel-Crafts, a mild and greener alternative to the classical analogue.

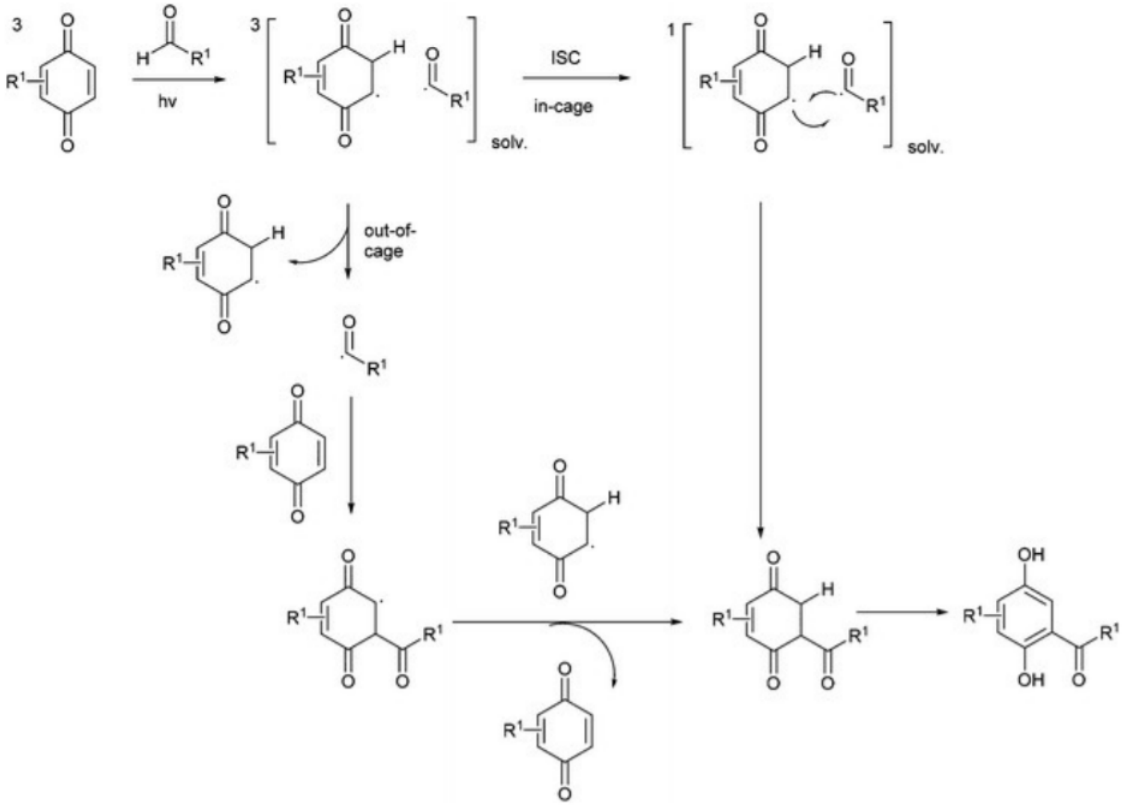


Figure 2.4.4.3.8: The solar-induced Friedel-Crafts reaction in which the elegance of the photoreaction is revealed by its simplicity, atom economy, and efficiency. Reproduced from Green Chemistry 2013, 15, 2830-2842. DOI:10.1039/C3GC41477A.

Ball Milling

Ball milling works on the principle of impact and attrition; size reduction results from impact as the balls drop from near the top of the shell. A ball mill is made up of a hollow cylindrical shell rotating about its axis. The axis may be either horizontal or at an acute angle to the horizontal. The shell is partially filled with balls whose grinding by action of balls, made of steel (chrome steel), stainless steel, ceramic, or rubber, results in much finer particles. The inner surface of the cylindrical shell is usually lined with an abrasion-resistant material such as manganese steel or rubber whereas less wear takes place in a rubber-lined mill. One of the reactions that is done in this manner is polymerization of MMA (methylmethacrylic acid) to PMMA (poly-) shown in 2.4.4.3.9. Below is a video resource that shows the process of ball milling. In this video it is possible to see how larger materials could be broken down and ground to specific sizes based on the size of grinding ball used.



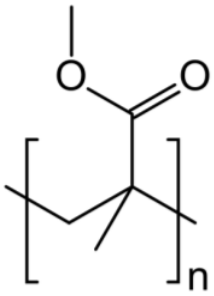


Figure 2.4.4.3.9: The opportunity to make one of the world's most abundant synthetic materials shown above (PMMA) through ball milling addresses fundamental green chemistry principles.
https://www.wikiwand.com/en/Poly(methyl_methacrylate)

The polymerization of MMA is normally done through one of several chemical-initiated approaches. In the ball milling approach, the application of mechanical energy is sufficient to conduct the reaction, a significant finding! The same type of approach can be used to extract lignin, the third or fourth most abundant polymer on the planet. A generic molecular representation of the structure of lignin is shown in 2.4.4.3.10.

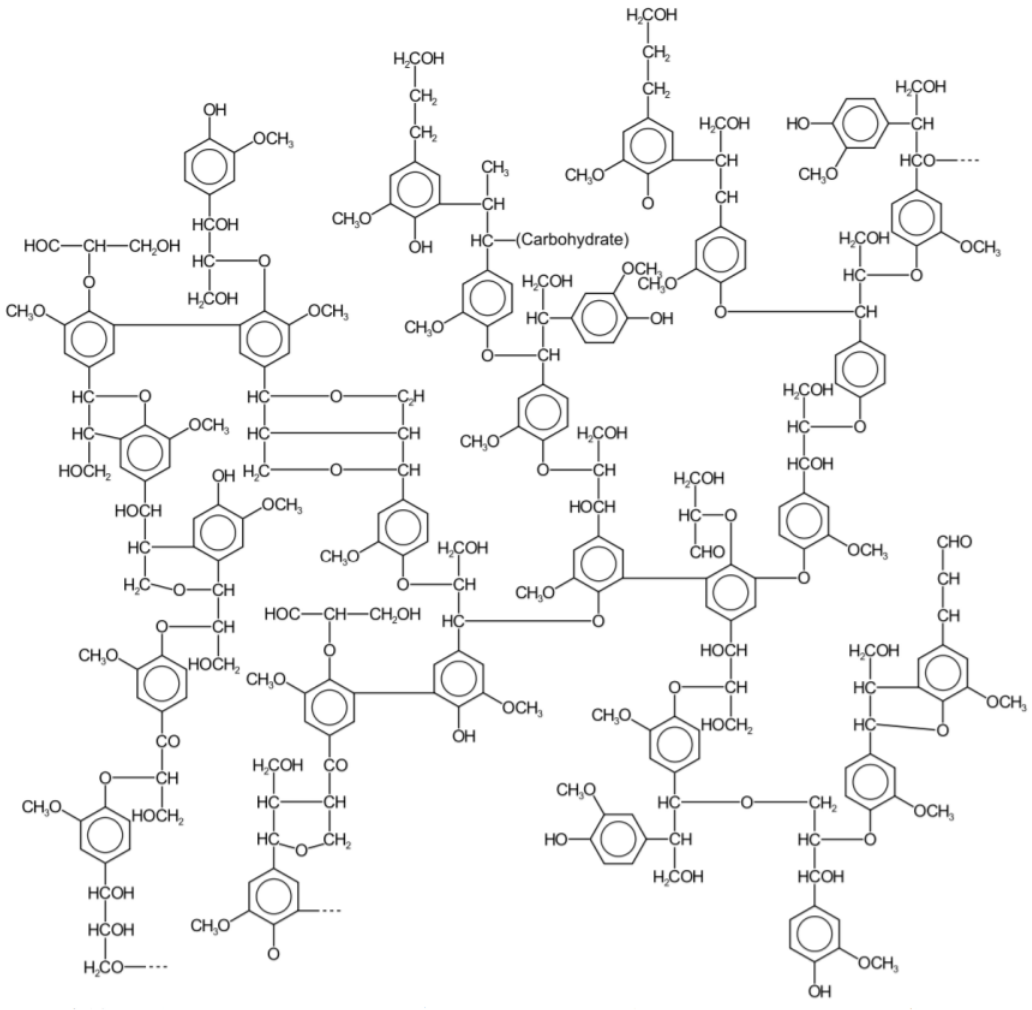


Figure 2.4.4.3.10: A chemical representation of the two-dimensional structure of lignin as it is known from gymnosperms (softwoods). https://en.Wikipedia.org/wiki/Lignin

Process Intensification

Process intensification can be defined as a strategy for introducing dramatic reductions in the footprint of a chemical plant to reach a given production objective. These reductions may consist of shrinking the size of pieces of equipment and reducing the number of unit operations or apparatuses. Such reductions tend to be significant because the objective is to dramatically reduce energy, materials waste, and process efficiency. Several of its most salient characteristics are:

- http://tinyurl.com/bpwpah9
- Continuous, short contact times
- Minimizes further reaction
- Higher, purer yields
- Mixing & heat transfer are very good! No explosive limits reached!

Reactive Extrusion

Reactive extrusion is a chemical engineering process characterized by the forced mixing of one or more components under high pressure conditions for a specific end goal. Hot melt extrusion is one example of this. It is defined as the application of heat and/or pressure to melt a polymer and force it though a small space (extruder) as part of a continuous process. It is a well-known process that was developed to make polymer products of uniform shape and density. It is widely applied in the plastic, rubber and food industries to prepare more than half of all plastic products including bags, films, sheets, tubes, fibers, foams, and pipes.

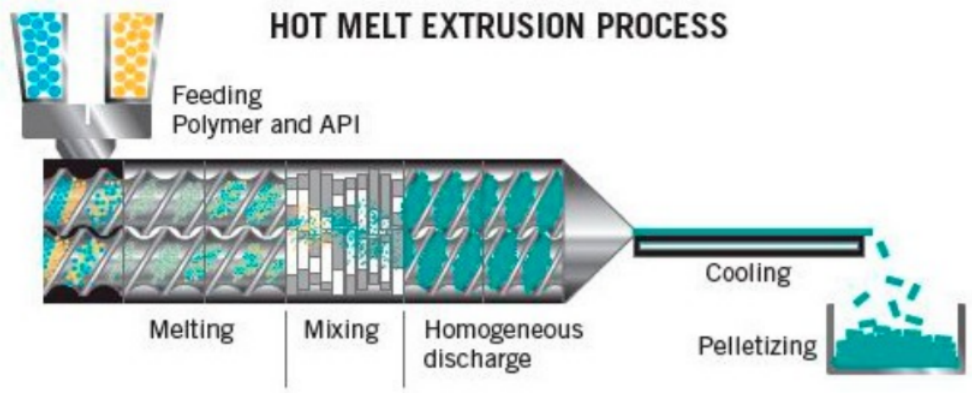


Figure 2.4.4.3.11 +*A simple idealization of the hot melt extrusion process in which the feed (polymer and API or Anionic Polymerization Initiator) go into a melting and mixing stage to ultimately yield pellets upon cooling. www.particlesciences.com/news...extrusion.html*

Supercritical Fluid

A supercritical fluid (sCF) is a substance at a temperature and pressure that are above the critical point, at which distinct liquid and gas phases do not exist. It is a very unique phase that can effuse through solids like a gas, and dissolve materials like a liquid.

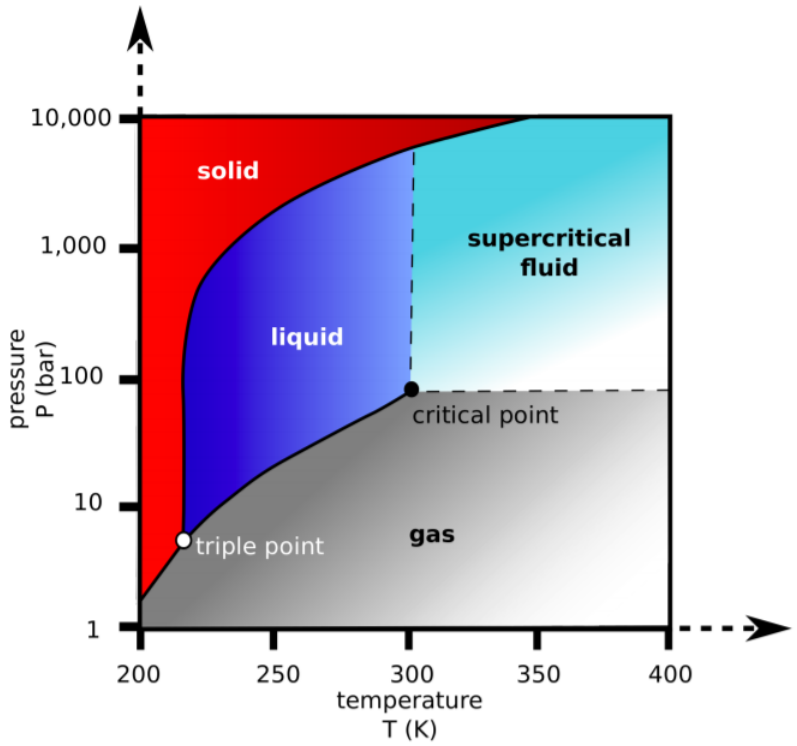


Figure 2.4.4.3.12: A phase diagram for the various states a substance can exhibit as a function of pressure and temperature. Notice that the upper right “quadrant” is the so-called supercritical fluid state. en.Wikipedia.org/wiki/Supercritical_fluid

sCFs are extremely useful in green chemistry because they can be derived from environmentally friendly materials such as water and carbon dioxide with little to no impact on the carbon footprint of the planet.

Water as a Solvent

Water can behave as an exquisite solvent in a host of typical organic reactions by virtue of its ability to encourage reactivity via the “hydrophobic effect”.

Microemulsions

Microemulsions are clear, thermodynamically stable, isotropic liquid mixtures of oil, water and a surfactant that are frequently in combination with a co-surfactant. The aqueous phase likely contains salt(s) or other ingredients, whereas the “oil” may actually be

a complex mixture of different hydrocarbons and olefins.

This page titled 2.4.4.3: Solvent Alternatives is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **4.3: Solvent Alternatives** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.4.4: Conclusions and Review

The opportunity to use alternatives to typical organic solvents to do the same transformations with the same or better efficiencies has never been better than today. We are learning that we can do a lot of chemistry in a green way because we are putting our efforts into it. Although we still have a long way to go before solventless, gas phase, etc., are the mode of doing chemistry, the range of available reactions and processes is superb.

Review Questions

1. What are the implications of using alternative solvents from a life cycle analysis perspective?
2. What are the energetic considerations for allowing polymerization to occur by ball milling?
3. What is the difference between a “critical point” and the “triple point” as defined in standard phase diagrams?

Further Reading

1. Cagniard de la Tour C. 1822 Exposé de quelques résultats obtenu par l’action combinée de la chaleur et de la compression sur certains liquides, tels que l’eau, l’alcool, l’éther sulfurique et l’essence de pétrole rectifiée. Ann. Chim.
2. Phys. 21, 127–132.
3. Andrews T. 1869 The Bakerian Lecture—On the continuity of the gaseous and liquid states of matter. Phil. Trans. R. Soc. Lond. 159, 575–590.
4. Licence P, Ke J, Sokolova M, Ross SK, Poliakoff M. 2003 Chemical reactions in supercritical carbon dioxide: from laboratory to commercial plant. Green Chem. 5, 99–104.
5. Tang SLY, Smith RL, Poliakoff M. 2005 Principles of green chemistry: PRODUCTIVELY. Green Chem. 7,761–762.
6. Lovelock KRJ, Villar-Garcia IJ, Maier F, Steinrück HP, Licence P. 2010 Photoelectron spectroscopy of ionic liquid-based interfaces. Chem. Rev. 110, 5158–5190.
7. Ball Milling Reference: "Ball Mill Critical Speed & Working Principle". YouTube. N.p., 2016. Web. 29 Nov. 2016.

This page titled 2.4.4.4: Conclusions and Review is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **4.4: Conclusions and Review** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

SECTION OVERVIEW

2.4.5: Alternative Reagants

2.4.5.1: Introduction to Reagants

2.4.5.2: Reagant Concepts

2.4.5.3: Specific Replacements

2.4.5.4: Conclusions and Review

This page titled 2.4.5: Alternative Reagants is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

2.4.5.1: Introduction to Reagants

Reagants

“A reagent ain’t supposed to make you faint!” -Lucian Lucia

Reagents, not unlike solvents in the previous chapter, are the chemical materials or reactants used in a chemical conversion to product. When we think of a “reagent”, we tend to think of a chemical used in a reaction, although there are numerous examples of chemicals throughout our world that do not neatly fit that mold. Dirt, water, air, etc., can all be reagents in terms of transformations to products. For example, the humic and fulvic acids found in dirt can act as exquisite reagents in separation, uptake, and cellular metabolic processes. Fulvics display a sorptive interaction with environmental chemicals before or after they reach concentrations that are toxic. The toxic herbicide Paraquat is rapidly detoxified, as are other organic compounds applied to the soil as pesticides. It is also vital to aid in the formation of new species of metal ions that bind with pesticides and herbicides to catalyze their breakdown.

We must think of a reagent as a means to an end. How can we use ingenious chemistry to take an otherwise invaluable or low-value item (commodity) to a valuable end product? This is the secret to enhancing the appeal and utility of green chemistry.

This page titled 2.4.5.1: Introduction to Reagants is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **5.1: Introduction to Reagants** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.5.2: Reagant Concepts

Major Concepts

The opportunity to do chemistry by means of, for example, a fixed substrate is very much in keeping with the concepts that we are learning. We explored the functionality and power of packed-bed columns that do chemistry normally done in solution with dissolved solutes. A packed-bed reactor can provide a source of protons, ions, or other chemicals that can affect chemistry. An ion-exchange system can perform ion metathesis, i.e., substitute the anion (X^{-m}) in dilute salt system $(M^{+n})_m(X^{-m})_n$ for the anion enriched on the packed bed. Shown below is an example of the process of interest. The solute (a mixture of proteins) is loaded onto the exchanger and subsequently adsorbs to the carboxymethyl anion substrate (a porogen). This process is typically characterized as a facile equilibrium process in which the protein mixture can replace the sodium cations. UV absorbance can follow the overall displacement as shown in the graph at the lower right (1^{st}: Mes anion peak; 2^{nd} and 3^{rd} peaks are the desired protein chloride substance).

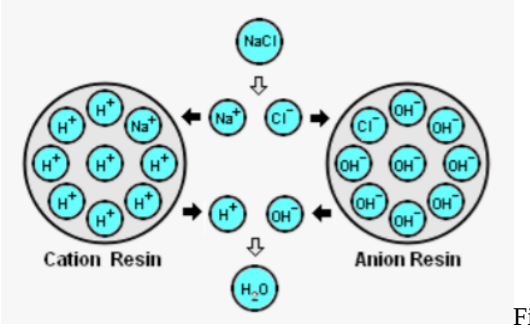


Figure 2.4.5.2.1: Shown is a simplified representation of the way metathesis occurs on cation and anion exchanging beds. The steps of the process are replacing the sodium cations for the positively charged hydrogens while simultaneously the chloride is replaced by the negatively charged hydroxides to lead to water.
http://www.pharmaguideline.com/2014/09/mixed-bed-ion-exchanger-inpurified-water- system.html

Replacing Hazardous Substances

A common misconception in chemistry is that we should be vigilant only when we know we are working with a hazardous compound. In fact, all substances should be treated with the level of respect that they deserve. Even salt can be an irritant when used incorrectly. The most importance criteria to consider in the evaluation of the hazards of substances are the following:

- **Efficacy:** the alternative must carry out desired transformation with comparable or superior efficiency
- **Safety:** the alternative reagent should display reduced volatility, flammability, toxicity, and/or reactivity, as well as increased stability
- **Environmental impacts:** the alternative reagent should represent a reduced environmental impact

Each of the above three criteria may be used in the assessment of a compound with respect to hazard level. Obviously, efficacy must be emphasized or else there is no need to go any further with the replacement.

Inorganic and Organic Supports

Supports or scaffolds or templates represent a fixed medium for the express purpose of a targeted transformation. We will use the following paper (shown are title, authors, and abstract) for our discussion in this part of the chapter.

“Paper-immobilized enzyme as a green microstructured catalyst” Hirotaka Koga, Takuya Kitaokabc and Akira Isogaia; DOI: 10.1039/c2jm30759f

The facile and direct introduction of methacryloxy groups into cellulose paper was carried out using a silane coupling technique, leading to the improvement of hydrophobicity and both dry and wet physical strengths of the paper. Immobilization of lipase enzymes onto the methacrylate- modified paper was then accomplished, possibly due to hydrophobic interaction. The as-prepared immobilized lipase on methacrylatemodified paper possessed paper-specific practical utility. During a batch process for the nonaqueous transesterification between 1 phenylethanol and vinyl acetate to produce 1- phenylethylacetate, the paperimmobilized lipase showed high catalytic activity, selectivity and reusability, suggesting that the methacryloxy groups introduced into the cellulose paper played a key role in the hyperactivation of lipases. In addition, a higher productivity of 1-phenylethylacetate was

achieved in a continuous flow reaction system than in the batch system, indicating that the interconnected porous microstructure of the paper provided favorable flow paths for the reactant solution. Thus, the paper-immobilized enzyme is expected to offer a green catalytic material for the effective production of useful chemicals.

What we note in this "paper" is the "smart" exploitation of paper as a porous microreaction chamber (interconnected paths) for a transesterification reaction between 1-phenylethanol and vinyl acetate to produce 1-Phenylethylacetate, as shown below:



Figure 2.4.5.2.2: The reaction of interest (transesterification) on a solid support (cellulosic paper) between 1 phenylethanol and vinyl acetate that leads to the production of 1- phenylethylacetate. Reproduced from "Paper-immobilized enzyme as a green microstructured catalyst" Hirotaka Koga, Takuya Kitaokabc and Akira Isogaia; DOI: 10.1039/c2jm30759f

This research effort explores the opportunity to catalyze reactions of interest using natural catalysts (enzymes). In this case, lipase-catalyzed reactions (e.g., esterification, transesterification, aminolysis, acylation, and thio-transesterification) that can be carried out in nonaqueous media have been examined in the synthesis of many useful chemicals for food, cosmetic, pharmaceutical and biodiesel.

However, lipases are typically unstable in nonaqueous media due to aggregation, denaturation, or some other form of deactivation. Enzymes are often immobilized on supporting materials (silica, ceramics, carbon, resins) to provide reusability, isolation advantages, and stability in nonaqueous media. Their study offered a number of advantages for the use of this green system including the asymmetry of the final productcomposition:

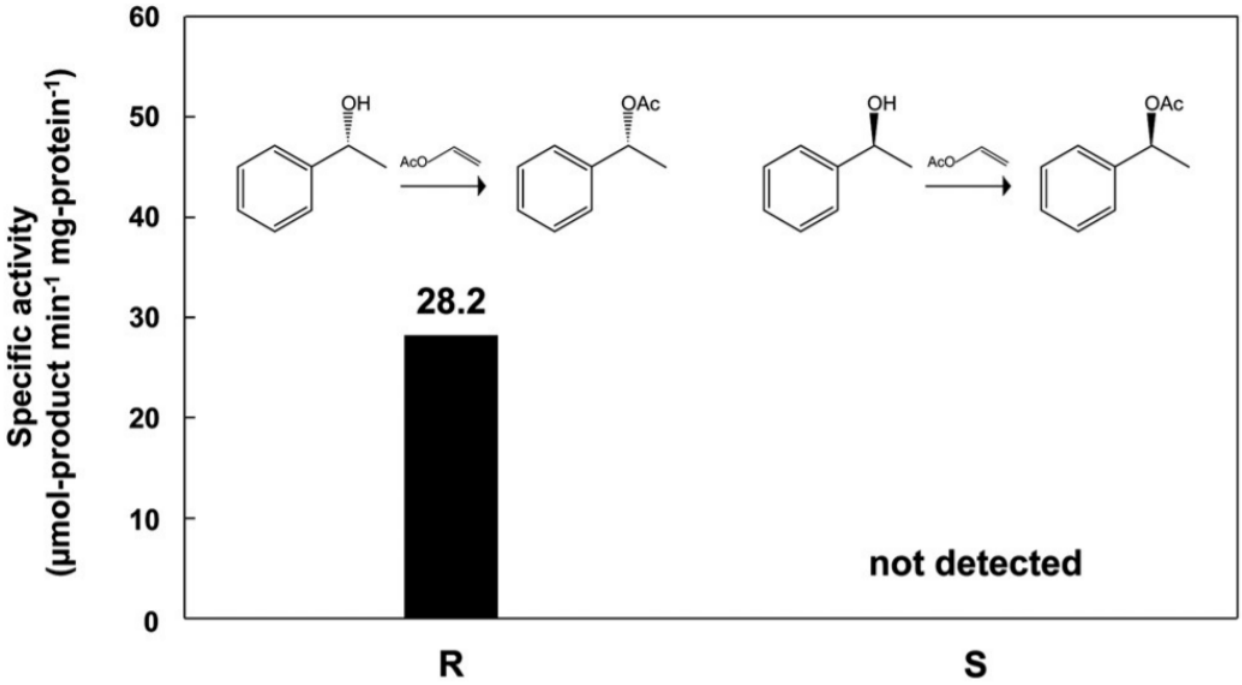


Figure 2.4.5.2.3: Enantioselectivity of immobilized lipase on methacrylate-modified Cellulose Whatman paper. Reaction temperature = 23 °C; stirring rate =150 rpm. Reproduced from "Paperimmobilized enzyme as a green microstructured catalyst" Hirotaka Koga, Takuya Kitaokabc and Akira Isogaia, DOI: 10.1039/c2jm30759f

Figure 5-3 shows that the paper-based system encourages the formation of the "R" enantiomer as opposed to the S (non-detected). Such a finding signals the value of a simple, facile, and highly conserved approach to an enantioselective and high yield (50%, 5 hrs.).

This page titled 2.4.5.2: Reagant Concepts is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **5.2: Reagant Concepts** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.5.3: Specific Replacements

Alternative Reagent Strategies

Removal of Lactose

The opportunity to remove lactose through a separation process can be achieved by attaching glutaraldehyde (two terminal-ended aldehyde group below, left) to an aminated silica particle (image on right) which is then covalently linked to a galactosidase enzyme that can be used to remove lactose from milk products followed by coupling with galactosidase and reaction with milk.

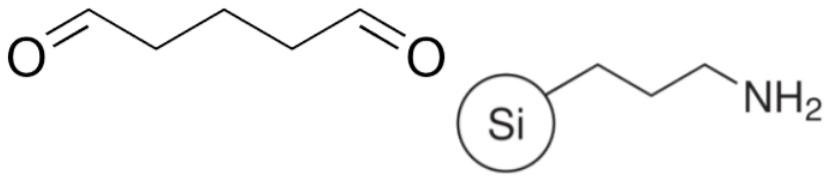


Figure 2.4.5.3.1: A depiction of two of the reagants used in the removal of lactose.

Quaternization of Amines

Quaternary ammonium cations, **quats**, are positively-charged ions of the structure shown below:

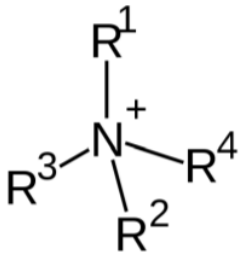


Figure 2.4.5.3.2: A simple representation of a quaternized amine or ammonium cation. https://www.wikiwand.com/en/Quaternary_ammonium_cation

The R group can be an alkyl group or an aryl group. Unlike ammonium (NH_4^+), quaternary ammonium cations are permanently charged, independent of pH.

Combinatorial Chemistry

Combinatorial chemistry is an approach to chemical synthesis that makes possible the preparation of a large number (e.g., millions) of compounds in a single sweep. These libraries can be mixtures, individual compounds, or computational structures.

Combinatorial chemistry can be used for small molecules, as well as for biomacromolecules such as peptides. Synthesis can quickly lead to large numbers of products. As an example, a molecule with three points of functionality or diversity (R_1, R_2, and R_3) can generate structures, where $N_{R\#}$ = the numbers of different substituents utilized. The basic principle of combinatorial chemistry is the preparation of large numbers of compounds that are then used to identify the useful components of the libraries.

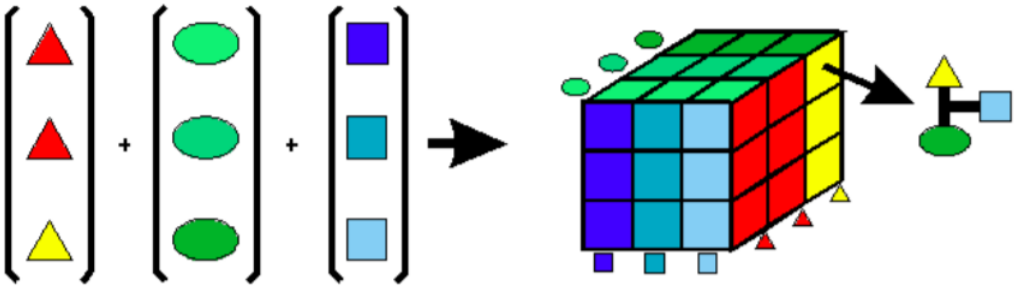


Figure 2.4.5.3.3: A simplified accounting of the process involved in the combinatorial chemistry approach toward developing a library of 3x3x3 indexed combinations. chem3513- 2007.pbworks.com/w/page/15648417/Combinatorial%20Chemistry

Although combinatorial chemistry has been in play over the last 25 years, its roots go to the 1960s when a Rockefeller researcher, Bruce Merrifield, started investigating solid phase synthesis of peptides. It has had its biggest impact in the drug industry where researchers optimize the activity profile of a compound by creating a library of many different but related compounds.

Starch

Starch is an energy storage biomaterial generated from carbon dioxide and water during photosynthesis. Among the panoply of natural polymers, starch is of considerable interest because of its biodegradability, low cost, renewability, and biocompatibility. It is therefore considered a promising candidate for the sustainable development of new functional materials.

Starch is composed of two homopolymers of D-glucose: amylase, a virtually linear α-D- (1, 4')- glucan, and branched amylopectin, with many α-1, 6'-linked branch points (Figure 2.4.5.3.4).

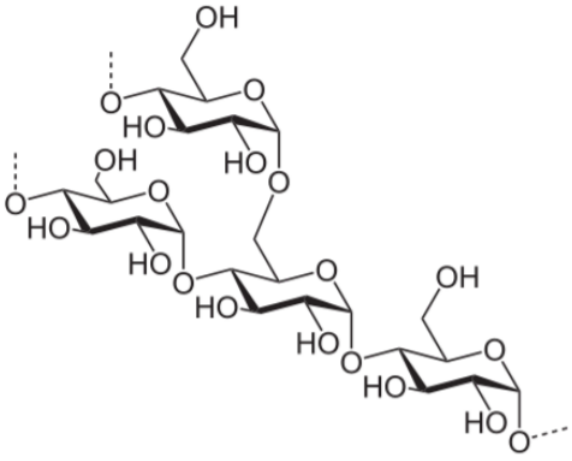


Figure 2.4.5.3.4: A simplified representation of the amylopectin macromolecule with the characteristic 1,6-linked branching. https://www.wikiwand.com/en/Amylopectin

Starch is a polyol: it has therefore, two secondary hydroxyl groups at C-2 and C-3, and one primary hydroxyl at C-6 when it is not linked. It is also quite hydrophilic and can be oxidized and reduced for the formation of ethers and esters and many other functional molecules. Starch has various proportions of amylose and amylopectin from about 10–20% amylose and 80–90% amylopectin based on the source. It forms a helical structure and occurs naturally as discrete granules because the short- branched amylopectin chains can form helical structures to crystallize. It has a MW of 500-20K consisting of a-(1à4)-D-glucose units whose extended shape has a 7-22 nm hydrodynamic radius range. The helical form of the molecule usually forms a stiff left- handed single helix whose structure, not unlike DNA, is supported by H-bonding between O2 and the O6 atoms. Amylose can undergo syneresis (dehydration synthesis) between vicinal (neighboring) glucan residues to form the cyclodextrin cavity in which the interior maintains a hydrophobicity. Interestingly, it can form double-stranded crystallites that are resistant to amylase (the putative enzyme for its deconstruction), and H-bonding and solubility. The crowing macromolecule from starch, CD, is a veritable green chemist's dream owing to a number of wonderful properties. Shown below is the toroidal structure of CD and dimensions:

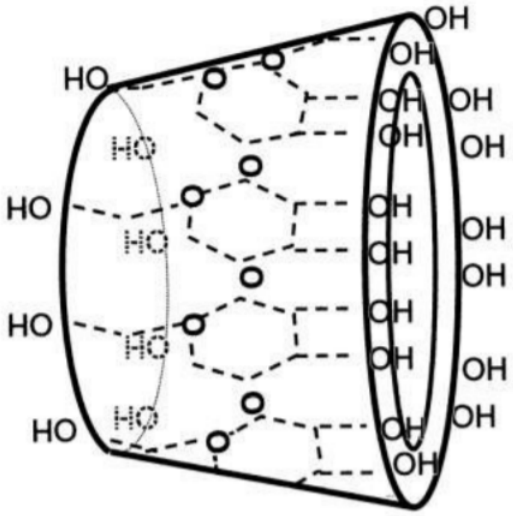


Figure 2.4.5.3.5: A representation of the CD macromolecule.
https://commons.wikimedia.org/wiki/F...cone_shape.jpg

How do Cyclodextrins Work?? https://www.youtube.com/watch?v=UaEes6cJu7k (How do Cyclodextrins work? (1980). Cyclolab.)

Specifically, it is a cyclic oligomer of a-D-glucopyranose held together by glycosidic bonds very much like what is seen in all polysaccharides. It was discovered in 1891 by Villiers who synthesized by the enzymatic conversion of amylose followed by a selective precipitation. The precipitation can be done with several solvents, albeit organic in nature and not sustainable, but with varying degrees of yields. Note that we look at three forms of the CDs as shown in the graph below:

	Precipitating Agent	Yield (%)
a-CD	1-decanol	40
b-CD	toluene	50-60
g-CD	Cyclohexadec-8-en-1-ol	40-50

What is remarkable about the CDs is their ability to form host/guest inclusion complexes owing to their unique HLB (Hydrophilic Lipophilic Balance) criteria. These are quite literally molecularscale reactors that can do many things. In the context of biomedical applications, the focus for our particular discussion, the following are among their benefits.

- To increase aqueous solubility of drugs
- To increase chemical stability of drugs
- To enhance drug delivery to and through biological membranes
- To increase physical stability of drugs
- To convert liquid drugs to microcrystalline powders
- To prevent drug-drug and drug-excipient interactions
- To reduce local irritation after topical or oral administration
- To prevent drug absorption into skin or after oral administration

The nanoreactors have an uncanny capacity to encapsulate and/or do chemistry with a number of chemicals that are hydrophobic or low polarity. Shown in Figure 5-9 is a cartoon that depicts the hydrolytic action of CD with a phosphate molecule.

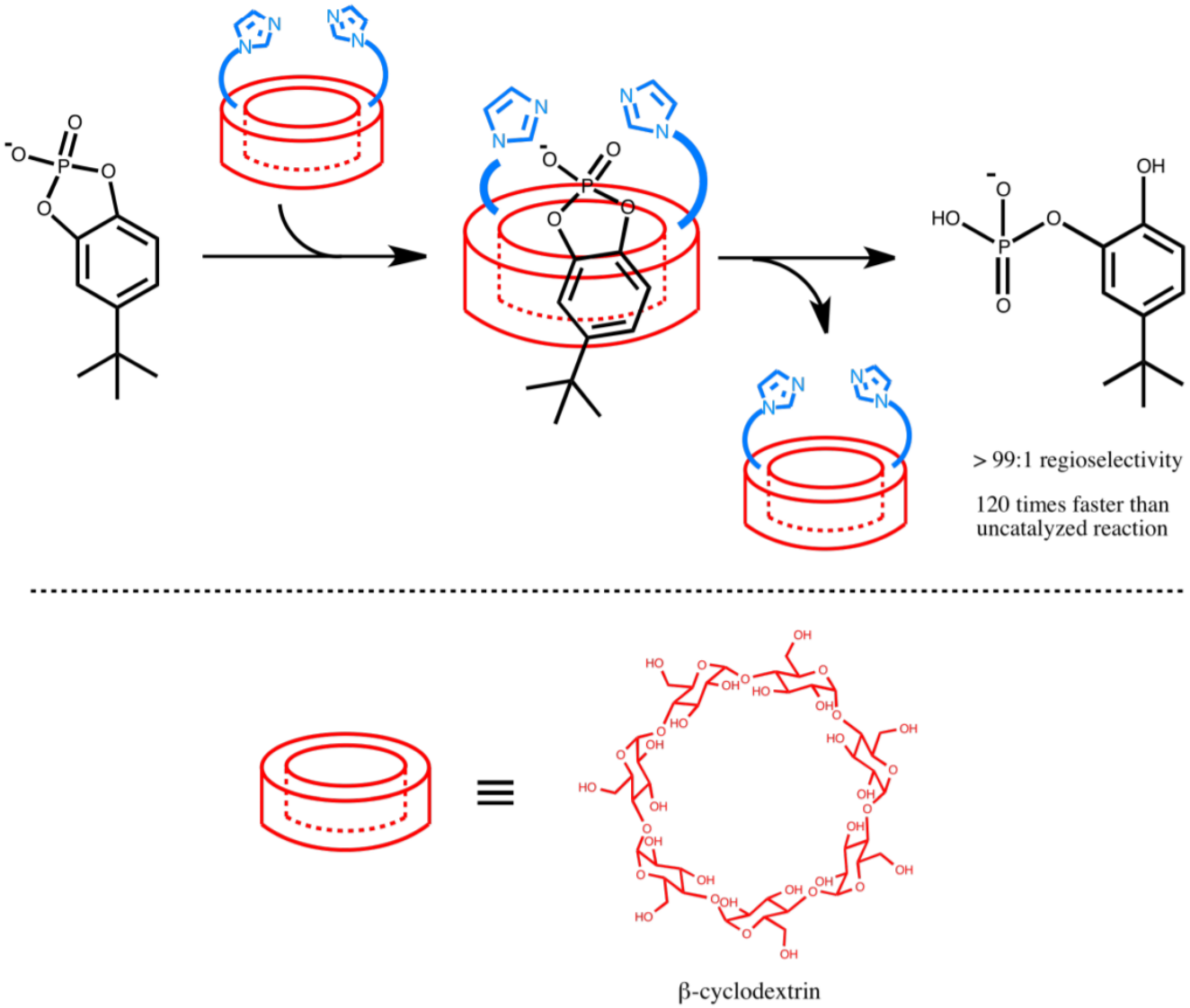


Figure 2.4.5.3.6: A simplified representation of the encapsulating and chemical reactivity behavior of a CD macromolecule for phosphate hydrolysis. https://commons.wikimedia.org/wiki/F...ease_Mimic.png

CD can by virtue of the varieties in its sizes accommodate relatively small to large guest molecules. Shown below in Figure 5-10 is a pictorial description of a 1:2 guest:host complex that engages in photophysical behavior.

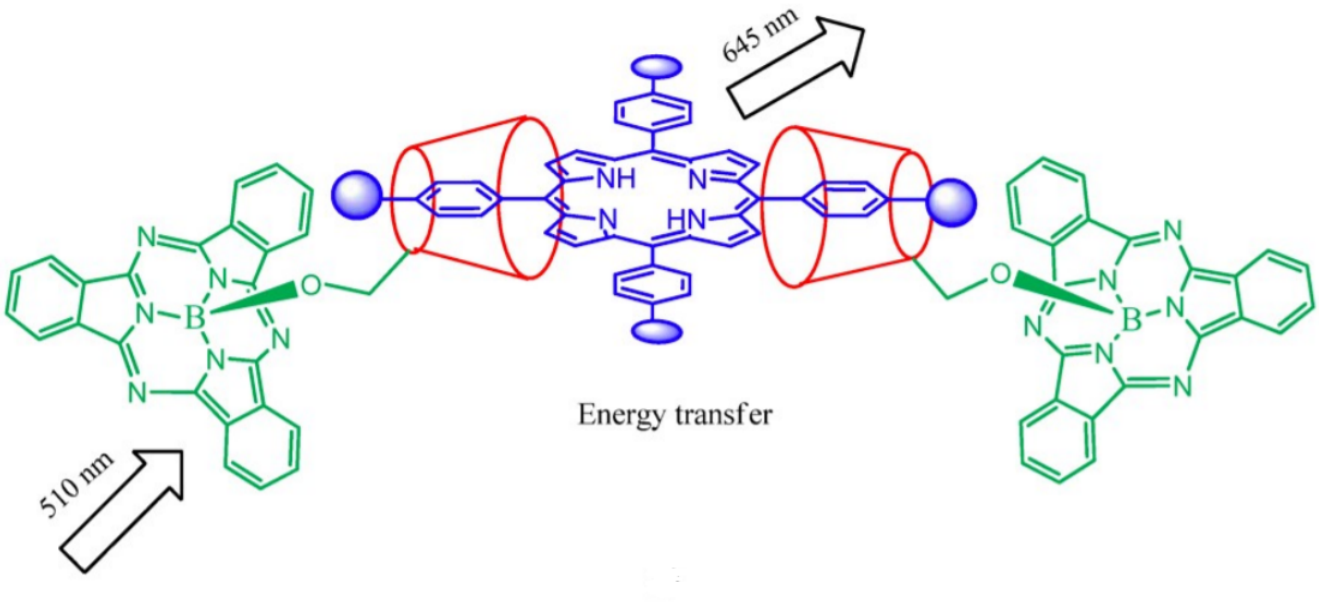


Figure 2.4.5.3.7: Pictorial description of a 2:1 host-guest complex formed by β-cyclodextrin- conjugated subphthalocyanine and tetrasulfonated porphyrin (4 outside blue spheres) as part of an energy transfer system. Reproduced from Molecules 2012, 17(10), 11763- 11799; doi:10.3390/molecules171011763.

The system is engaging in an energy transfer from the peripheral boron- containing molecules to the encapsulated porphyrin molecules which releases a photon of light upon resonant energy transfer (645 nm) from the peripheral molecules.

Phase Solubility Diagrams

Phase solubility studies are carried out in aqueous systems at different temperatures to calculate stability (equilibrium) constants, (K_c) and thermodynamic values for the formation of inclusion complexes. It is known that if phase solubility diagrams show that the solubility of a guest molecule increases linearly along with the concentration of CD, then they can be considered as AL-type phase diagrams [1], suggesting the formation of 1:1 complexes, which are the most common and best understood of these types of interactions. When the phase solubility data are collected at different temperatures, we can obtain valuable additional information such as the thermodynamic parameters for the formation of the complex.

The integrated form of the Van't Hoff Equation allows for the calculation of the enthalpy (DH) and entropy changes (DS) depending on the variations of the stability constants with temperature [2]:

$$ln(K_c) = \frac{D \times H}{R \times T} + \frac{D \times S}{R}$$

The K_c can also be expressed as the following relationship:

$$K_c = \frac{[D - CD]}{[D]} = [D] \times [CD]$$

that can be pictorially described as:

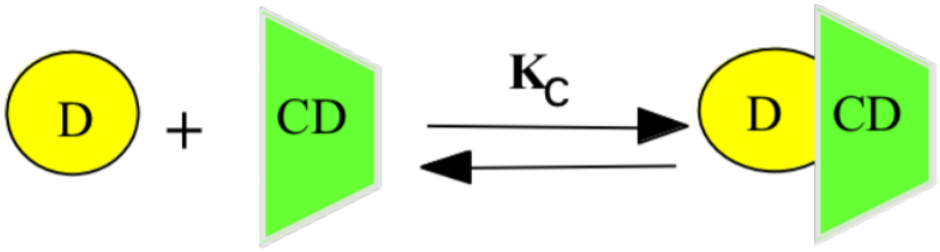


Figure 2.4.5.3.8: A picture describing the physical realization of the Van Hoff Equation

in which there is a complexation with D (the substrate or drug in this case) and the CD. The above represents an AL-type complex which is first order with respect to the CD and the drug (D).

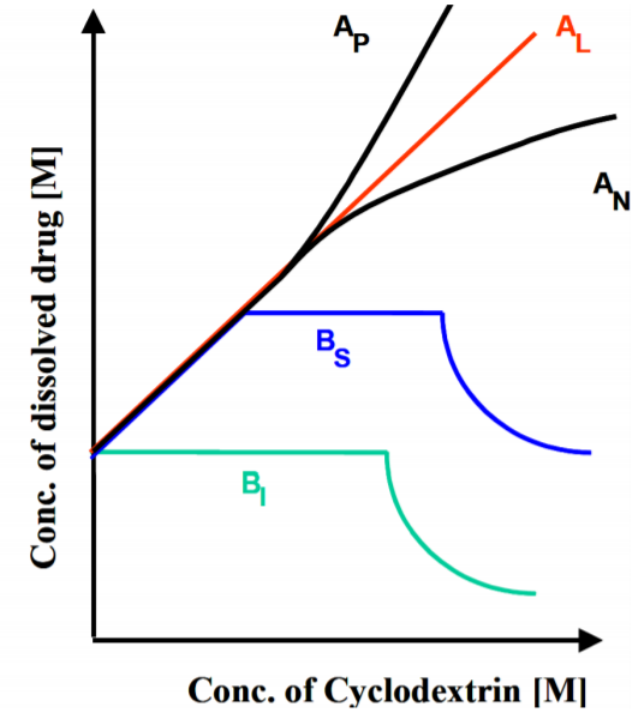


Figure 2.4.5.3.9: A mathematical profiling of the various phase solubility regimes available to substrate (drug) and ligand (CD) complexes as a function of CD concentration. Reproduced from Drug Discovery Today 2016, 21, 356-362. https://doi.org/10.1016/j.drudis.2015.11.017

A-type phase-solubility profiles are defined by the solubility of the substrate (drug) increasing with increasing ligand (cyclodextrin) concentration. When the complex is first order relative to the ligand and first or higher order with respect to substrate then A_L-type phase-solubility profiles are obtained. If the complex is first order with respect to the substrate but second or higher order with respect to ligand, A_P-types are obtained. A_N-type phase-solubility profiles are difficult to interpret. B-type phase- solubility profiles indicate formation of complexes with limited solubility in the aqueous medium. The water-soluble cyclodextrin derivatives form A-type phase- solubility profiles while less soluble cyclodextrins form B-type profiles. A number of drug/cyclodextrin complexes form inclusion complexes, butcyclodextrins form non- inclusion complexes and aggregates that dissolve drugs through micelle-like structures. The phase-solubility profiles do not verify inclusion complexes, but only describe how increasing cyclodextrin concentration influences drug solubility. As we already discussed, the most common type of cyclodextrin complexes are the 1:1 where one drug molecule (D) forms a complex with one cyclodextrin molecule (CD). In an AL-type phasesolubility diagram where the slope is less than unity, the stability constant ($K_{1:1}$) of the complex can be calculated from the slope and the intrinsic solubility (S0) of the drug in the aqueous media (i.e., drug solubility when no cyclodextrin is present):

$$K_{1:1} = \frac{Slope}{S \times (1 - Slope)}$$

https://bio.libretexts.org/@go/page/108733

The value of $K_{1:1}$ is typically 50 < 2000 M^{-1} with a mean value of 129, 490, and 355 M^{-1} for α-, β- and γ- cyclodextrin, respectively [3].

This page titled 2.4.5.3: Specific Replacements is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **5.3: Specific Replacements** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.5.4: Conclusions and Review

Advances in drug development such as high throughput have increased the total and variety of drug candidates whose clinical usefulness is principally limited to a great degree by the following:

- Water insolubility;
- Chemical or physical instability;
- Local irritation after administration

Cyclodextrins alleviate many of these issues. As evidence, there are currently around 30 different cyclodextrin pharmaceutical products on the market that include different types of pills, capsules, parenteral solutions, nasal sprays, suppositories, eye drops, and skin products. Cyclodextrins tend to replace organic solvents, enhance oral bioavailability, reduce gastrointestinal irritation, and increase dermal availability. In addition, animal and human studies have demonstrated that cyclodextrins improve drug delivery.

Addition of cyclodextrins to existing formulations will seldom result in acceptable outcome. We still lack deep knowledge of the physical factors involved in complex formation. Cyclodextrins form both inclusion and non-inclusion complexes in aqueous solutions. It has also been shown that cyclodextrins form aggregates that are able to solubilize in a micellar-like fashion; their exact structures are not known, nor is the way they influence drug delivery

Review Questions

1. What are ways that ion exchange columns can be improved to ensure maximum efficiency? In other words, what is the principle that governs their functionality?
2. Why is paper such a good medium for chemical reactions?
3. Does a practical limit exist for the variations of combinatorial chemical processes that can be studied?
4. Can you envision a way that CD systems can aggregate in human circulatory systems?

Further Reading

1. Cagniard de la Tour C. 1822 Exposé de quelques résultats obtenu par l'action combinée de la chaleur et de la compression sur certains liquides, tels que l'eau, l'alcool, l'éther sulfurique et l'essence de pétrole rectifiée. Ann. Chim. Phys. 21, 127– 132.
2. Andrews T. 1869 The Bakerian Lecture—On the continuity of the gaseous and liquid states of matter. Phil. Trans. R. Soc. Lond. 159, 575–590.
3. Licence P, Ke J, Sokolova M, Ross SK, Poliakoff M. 2003 Chemical reactions in supercritical carbon dioxide: from laboratory to commercial plant. Green Chem. 5, 99–104.
4. Tang SLY, Smith RL, Poliakoff M. 2005 Principles of green chemistry: PRODUCTIVELY. Green Chem. 7,761–762.
5. Lovelock KRJ, Villar-Garcia IJ, Maier F, Steinrück H-P, Licence P. 2010 Photoelectron spectroscopy of ionic liquid-based interfaces. Chem. Rev. 110, 5158–5190.

This page titled 2.4.5.4: Conclusions and Review is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **5.4: Conclusions and Review** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

SECTION OVERVIEW

2.4.6: Reaction Types, Design, and Efficiency

2.4.6.1: Introduction to Reactions

2.4.6.2: Reaction Concepts

2.4.6.3: Reaction Design Concepts

2.4.6.4: Conclusions and Review

This page titled 2.4.6: Reaction Types, Design, and Efficiency is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

2.4.6.1: Introduction to Reactions

Introduction

“What has one voice, and is four-footed in the morning, two-footed in the afternoon and three- footed at night”?

-The Sphinx (Oedipus)

In any reaction, a transformation from a starting material to a final product is the basis. In other words, a budding green chemist such as yourself will investigate how a chemical will undergo one or more treatments to produce a needed product in a manner consistent with the dogmas of GC. Typically, any chemist is chiefly interested in the concepts of yield and specificity, namely, how efficiently is the conversion done based on a theoretical yield and how exactingly is it done (with respect to a series of products). These latter terms have long been the “mantras” of organic, synthetic, and biochemical studies. The secret to any successful transformation is ensuring that we get high yield and maximum specificity no matter what else is generated. A successful chemical transformation for a green chemist, however, must go beyond these staid concepts and introduce the concepts of safety, waste, energy, and number of steps (which are reflected in the first three terms).

The history of chemistry has been one of neglecting such concepts because of the lack of environmental consciousness, low regard for safety, and lack of understanding of economics or life cycle. When we approach the riddle of doing chemistry, we have not done so with an eye to mimicking nature. We have always approached it anthropomorphically, that is to say, with the resolute conviction that we can do it better, but we fail.

This page titled 2.4.6.1: Introduction to Reactions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **6.1: Introduction to Reactions** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.6.2: Reaction Concepts

Overarching Concepts

Reaction types are significant to all fields of chemistry because they are what characterize any transformation from starting materials to products. In the pantheon of reactions, we tend to focus on functional group creation and atom transformation, both of which can be accomplished through the menu of reaction types shown below:

- Rearrangement (RAR)
- Addition
- Substitution
- Elimination
- Pericyclic reaction
- Ox/Red
- Reaction design
- Atom economy

The above six (6) general reaction categories simplify the panoply of reactivity paradigms in the discipline of chemistry. We can nearly list every pertinent reaction within one or more of the listings shown. Indeed, even acid/base chemistry can fall under Addition or Elimination reactions based on their precise reactivity patterns. We will delve deeply into each type in the following sections.

Rearrangements (RARs)

A *rearrangement reaction* covers a broad swath of chemistry which is principally characterized by the rearrangement of the skeletal configuration of a molecular system to yield a structural isomer (a swapped-out form of the original structure while retaining the original atoms). Oftentimes, a substituent moves from one backbone atom to another backbone atom. In the classical organic chemistry example shown below, we witness the Wagner-Meerwein RAR.

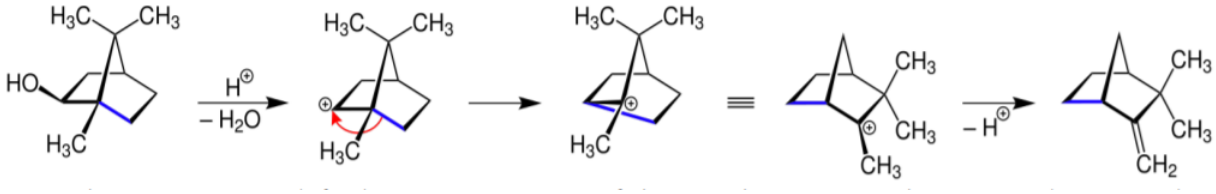


Figure 2.4.6.2.1: Shown is a simplified representation of the mechanistic pathway involving acid catalyzed transformation of the bicyclic tetra-substituted norbornane to an exo vinylic-substituted norbornane. The first step of the process is expulsion of water (dehydration) that leads to a reactive carbocation (albeit secondary) that lead to a 1,3- sigmatropic shift and delocalization of the carbocation to the bridgehead carbon (tertiary – hence, more stable). https://commons.wikimedia.org/wiki/F...hene_Mech1.svg

A rearrangement is not exactly well depicted by discrete electron transfers shown by curly arrows. The actual mechanism of alkyl groups moving in Figure 2.4.6.2.1 involves a fluid transfer of the alkyl group along a bond that is not described by simple ionic bond-breaking and reformation. Explanation by orbital interactions provides a better approximation although it is possible to draw the curved arrows for a sequence of discrete electron transfers to give the same result as a RAR reaction.

There are other types of RARs besides the chemically induced one shown above: we can have thermal or photo-induced as well. In these latter cases, heat or light can act as the triggers to actuate the transformation.

Thermal RARs

The isomerization of unsubstituted azulene to naphthalene was the first reported and hence most studied thermal transformation for an aromatic hydrocarbon. Many mechanisms have been suggested, yet not a one has been unequivocally confirmed.

Five mechanisms were proposed: reversible ring-closure, shown above, a norcaradiene- vinylidene, diradical, methylene walk, and a spiran mechanism. The reversible ring- closure mechanism is inaccurate despite its overall scientific appeal and evidence, and as such, multiple reaction pathways were deemed occurring simultaneously. This idea was widely accepted. It was thought that at high

temperatures one mechanism would be energetically favored although energetic studies displayed similar activation energies for all mechanisms [1].

Four mechanisms for thermal isomerizations are warranted: dyotropic, diradical, and two benzene ring contractions. A 1,2-carbon shift to a carbene preceding a 1,2-hydrogen shift, and a 1-2- hydrogen shift to a carbene followed by a 1,2-carbon shift. The dyotropic mechanism shows the 1,2-shift displayed in Figure 2.4.6.2.2 **.**

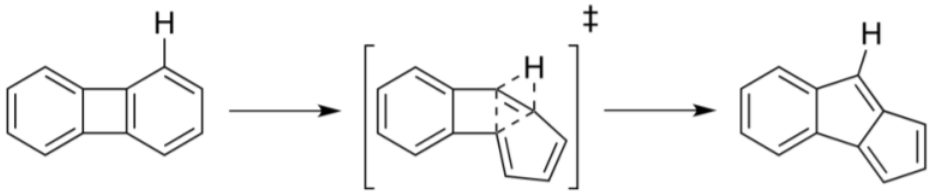


Figure 2.4.6.2.2 *Shown is a simplified representation of the mechanistic pathway involving the thermal transformation of the above tricyclic aromatic.* http://www.wikiwand.com/en/Thermal_rearrangement_of_aromatic_hydrocarbons

Photoinduced RARs

The mechanisms of organic photochemical reactions are a treasure trove for photochemists and physicists because they provide tremendous insight into fundamental electronic behavior as a function of energy-matter interactions. The absorption of ultraviolet light by organics may lead to a number of reactions. Over the last century, an immense quantity of photochemical reactions have become unearthed; it has been found that reactions that are unlikely in ground-states are accessible in electronic excited-state configurations. One of the earliest photochemical studies was on the natural product santonin. Ciamician (1857-1922) observed that under sunlight exposure, santonin gave several photoproducts. The structure of santonin (molecule on left in Figure 6-3) was first described by Clemo and Hayworth in 1929 whose initial photoproduct is lumisantonin. As depicted, the photoreaction involves a C-3 carbonyl group movement to C-2, the C-4 methyl has moved to C-1, and the C-10 carbon has been inverted.

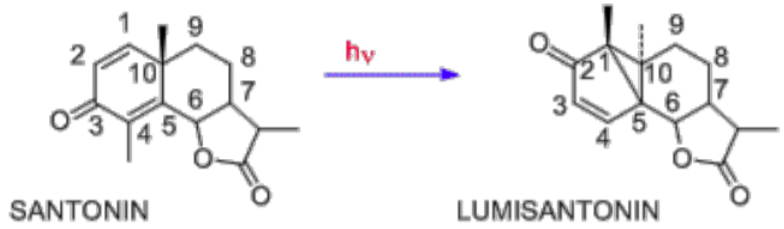


Figure 2.4.6.2.3: The photochemical transformation of santonin to lumisantonin as first discovered by Ciamician is represented above. https://en.Wikipedia.org/wiki/Mechan...photochemistry

Beckmann RAR

This reaction is named after the German chemist Ernst Otto Beckmann (1853–1923). It is representative of an acid-catalyzed rearrangement of an oxime to an amide (when the oximes are cyclic, they yield “lactams”).

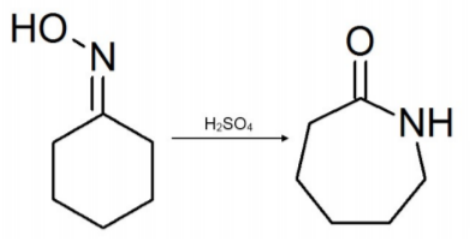


Figure 2.4.6.2.4: The classical transformation of a cyclic oxime to a lactam via the Beckmann RAR. http://www.wikiwand.com/en/Beckmann_rearrangement

Caprolactam, in this example, is a very important chemical because it is the feedstock in the production of Nylon 6, a material used extensively in the manufacture of textiles, clothes, carpets, cushions, bulletproof vests, etc. The reaction includes acetic acid, hydrochloric acid, sulfuric acid, and acetic anhydride, to catalyze the rearrangement. Sulfuric acid is the most commonly used acid for commercial lactam production because it forms ammonium sulfate (a common agricultural fertilizer) as a by-product when neutralized with ammonia.

In general, rearrangements tend to have no waste of atoms because they are a type of isomerization. They are hence very atom economical and hence very efficient modes of chemical transformations. At this point, we will delve into a series of rearrangements to further illustrate the point.

Addition Reactions

This type of reaction occurs when in general two molecules covalently bind to form a larger one referred to as the adduct. These reactions are limited to chemical compounds that have multiple bonds, such as molecules with carbon–carbon double bonds (alkenes), triple bonds (alkynes) hetero double bonds like carbonyl (C=O) groups, or imine (C=N) groups. An addition reaction is typically the reverse of an elimination reaction. For example, the hydration of an alkene to an alcohol is the reverse of dehydration which leads to an oxidative product (alkene). Electrophilic and nucleophilic additions are the main types of polar addition. The non-polar additions can be free radical and cycloaddition.

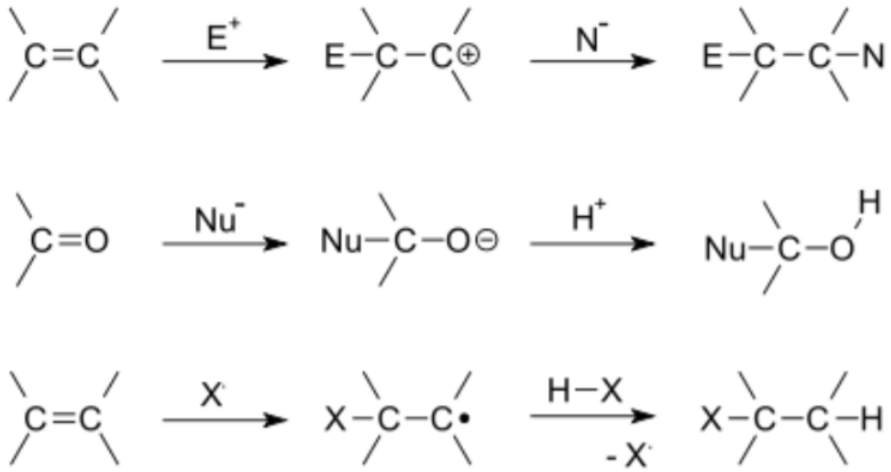


Figure 2.4.6.2.5: A generic depiction of the three major schemes of addition reactions. The top represents an electrophilic addition reaction (as shown in electrophilic aromatic addition reaction such as nitration, sulfonation, etc., catalyzed by Lewis acids); the middle represents a nucleophilic addition reaction (as represented by S_N2 reactions); the bottom represents a radical-induced addition reaction which as opposed to the top two, requires an activation step (development of a radical from the olefin) that then allows for an addition (in the second step). https://www.wikiwand.com/en/Addition_reaction

In general, addition reactions are very common because they are simple, very powerful, robust, and preserve/conserve atom economy during the reaction sequence (think of the mathematical operation of addition – all numbers sum!). Thus, in general, no additional by-products are generated allowing for a clean and straight forward reaction scheme that in the final analysis may require little to no purification steps.

Substitution Reactions

These types of reactions are also known as single displacement reaction or single substitution reaction). One functional group is replaced by a more reactive functional group or else it would not happen (the reverse would be faster or more thermodynamically stable). Substitution reactions are classified either electrophilic or nucleophilic depending upon the reagents. A good example of a substitution reaction is halogenation. When chlorine gas (Cl-Cl) is irradiated with the appropriate light energy, several of the molecules photolyze into two chlorine radicals (Cl·) whose electrons are very nucleophilic. In Figure 6-6, one of radicals can rupture (homolytically) a C-H bond through abstraction of one of the equivalent protons on methane to form electrically neutral H-Cl. The other radical reforms a covalent bond with the CH_3 . to form CH_3Cl.

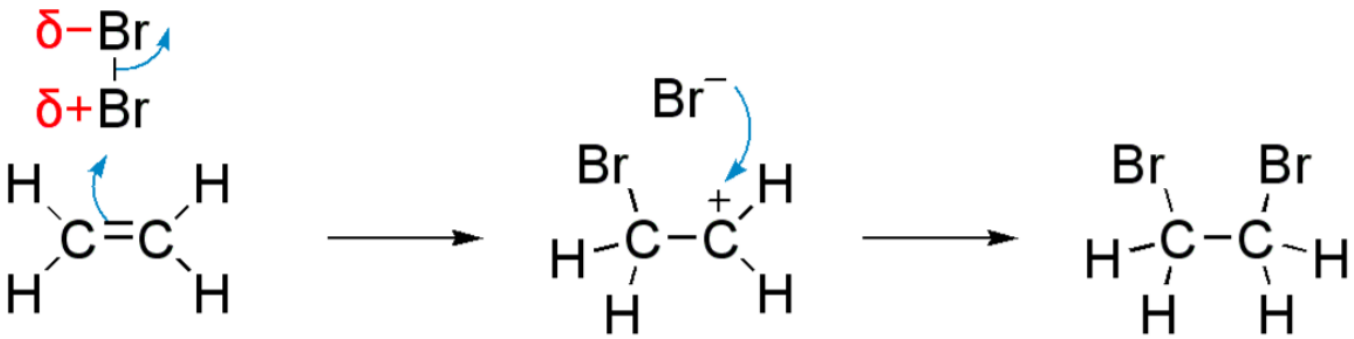


Figure 2.4.6.2.6 *An electrophilic substitution reaction is shown in which chlorine electrophilically attacks a double bond that results in the activation of two C-H bonds. commons.wikimedia.org/wiki/File:Electrophilic_reaction_of_bromine_with_ethene.png*

Elimination Reactions

An elimination reaction typically involves the removal of two substituents a substrate in either a one or two-step mechanism. The one-step mechanism is the **E2 reaction**, and the two-step mechanism is the **E1 reaction**. The numbering scheme relates to the kinetics of the reaction, i.e., bimolecular and unimolecular, respectively.

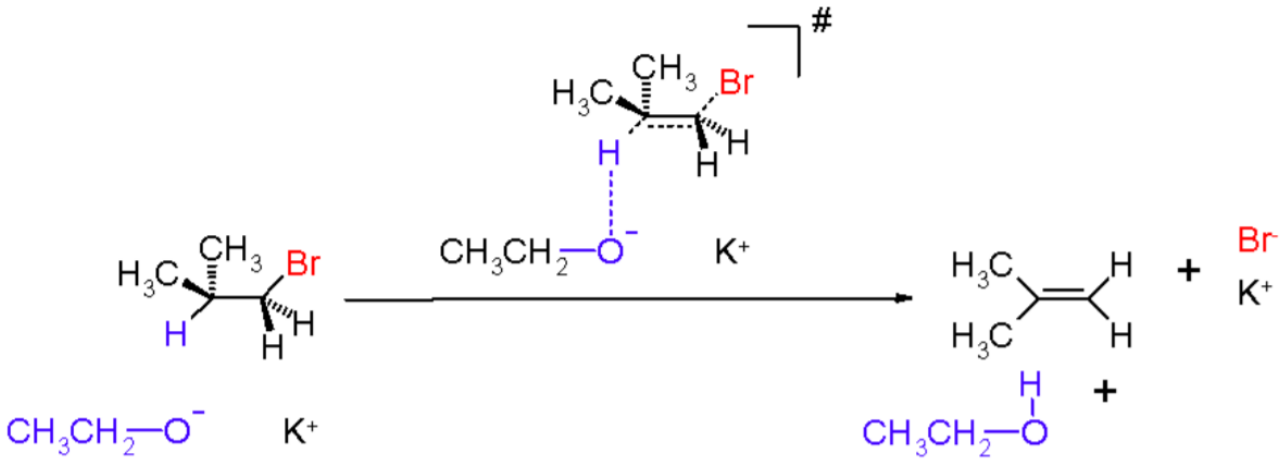


Figure 2.4.6.2.7: A representation of an E2 reaction (bimolecular) in which ethoxide acts as a base in a 1,2-atom removal to provide an olefin. commons.wikimedia.org/wiki/File:E2_Elimination_Reaction.png

Elimination reactions are important because they are among a very few set of useful reactions that generate unsaturation within organic molecules. They are the opposite of addition reactions and involve homolytic or heterolytic dissociation of molecular groups on adjacent (vicinal) atoms leading to increased bond order on these vicinal atoms.

For example,RCH_2CH_2OH → R-CH=CH_2 +H_2O.This latter reaction requires the expulsion of a water to lead to one degree of unsaturation. The presence of a leaving group is a pre-requisite for these types of reactions.

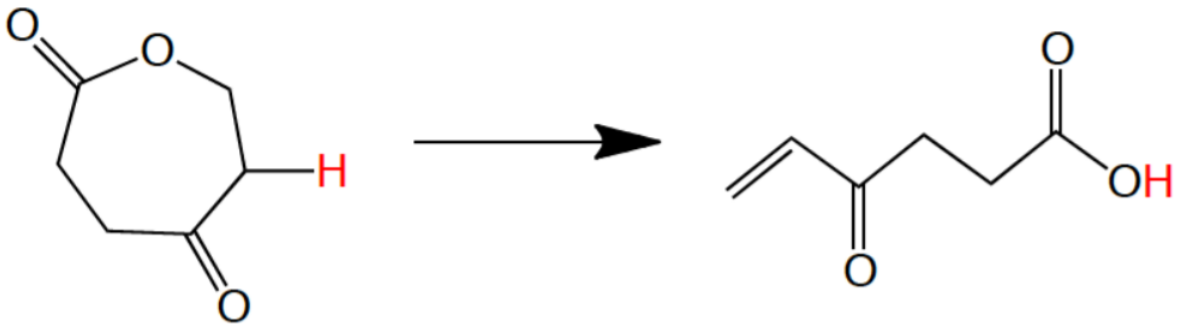


Figure 2.4.6.2.8: The 100% atom economic transformation of a lactone to an a,b-unsaturated carbonyl carboxylic acid in which the proton in red is conserved.

In Figure 2.4.6.2.8, the opportunity to retain a hydrogen on the final structure is an unusual mechanistic phenomenon. The conversion relies on a strict orbitally conserved concerted mechanism in which the degree of unsaturation (terminal vinyl group) is generated in tandem with the abstraction of the proton by the lactonic ring oxygen. Such a sigmatropic shift would be over 3 atoms (hence, an unusual 1,3-hydride shift).

Pericyclic Reactions

A pericyclic reaction (Figure 2.4.6.2.9) is a reaction in which the transition state of the molecule has a cyclic geometry and progresses in a concerted manner. They are typically rearrangement reactions. The major classes of pericyclic reactions are electrocyclic, cycloaddition, sigmatropic, group transfer, cheletropic, and dyotropic, all of which are considered to be equilibrium processes.

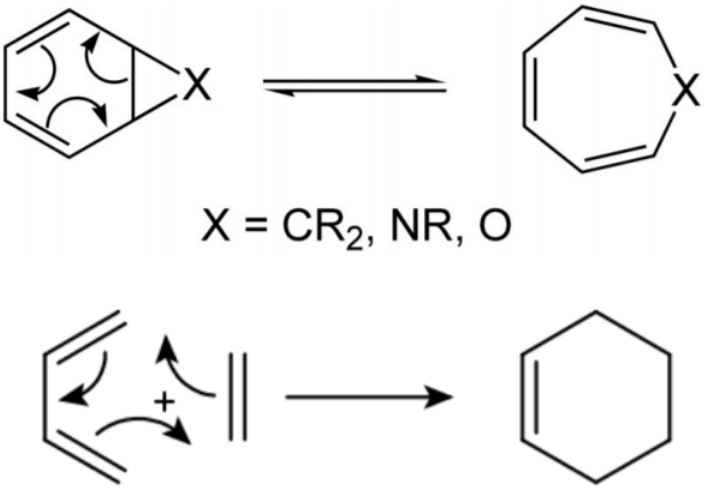


Figure 2.4.6.2.9: Classic examples of a pericyclic reaction that is bimolecular in nature as opposed to the preponderance of entropically-favored unimolecular pericyclic reactions. The above reactions are the classical [4+2] cycloaddition known universally as the **Diels-Alder** Reaction. commons.wikimedia.org/wiki/File:Diels_Alder_Mechanismus.svg

These types of reactions are governed by Frontier Molecular Orbital (FMO) Theory. Frontier molecular orbitals for a molecule are so called because the orbitals are the “frontier” or vanguard (leaders) of electron occupation. These are known classically as the highest-energy occupied (HOMO) and lowest-energy unoccupied (LUMO) molecular orbitals. The HOMO is characterized by nucleophilic or electron donation, whereas the LUMO is the opposite. Chemical reactions and resonance are very successfully described by the overlap between a filled HOMO and an empty LUMO.

Acidity and basicity are terms used in FMO Theory very broadly. Acidity refers to a ligand, metal center, or orbital ability to accept electron density (from electron sources that include Brønsted bases). Basicity refers to a ligand, metal center, or orbital ability to donate electrons (to electron sinks that include Brønsted acids). The theory distinguishes among σ-acids, σ-bases, π-acids, and π-

bases. The first two, σ-X, accept or donate electrons in a σ manner, aligned head on with another orbital, whereas the latter two, π-X, accept or donate electrons in a π manner, aligned side by side with another orbital.

A pericyclic reaction typically is unimolecular and zero order with respect to kinetics. As shown below, the bicyclic on the left is highly strained because of the 1,2 bridging and undergoes an electrocyclic RAR to relieve the steric stress associated with the heterocyclic cyclopropanoid moiety.

Diels-Alder

This is a [4+2] cycloaddition as already mentioned that occurs between a conjugated diene and a substituted alkene known as a dienophile to form a substituted cyclohexene. It was first described by Otto Diels and Kurt Adler in 1928 that lead to the Nobel Prize in Chemistry in 1950. It is a reliable method for forming 6- membered systems with good regio- and stereochemical control. The concept has been applied to other π-systems to provide the corresponding heterocycles (hetero-Diels–Alder reaction).

1,3-Dipolar Cycloaddition

This reaction occurs between a 1,3-dipole and a dipolarophile to form a five-membered ring (Figure 2.4.6.2.10). The reaction is sometimes referred to as the Huisgen cycloaddition which specifically describes the 1,3-dipolar cycloaddition between an organic azide and an alkyne to generate 1,2,3-triazole. Currently, this reaction is an important route to the region- and stereospecific synthesis of fivemembered heterocycles.

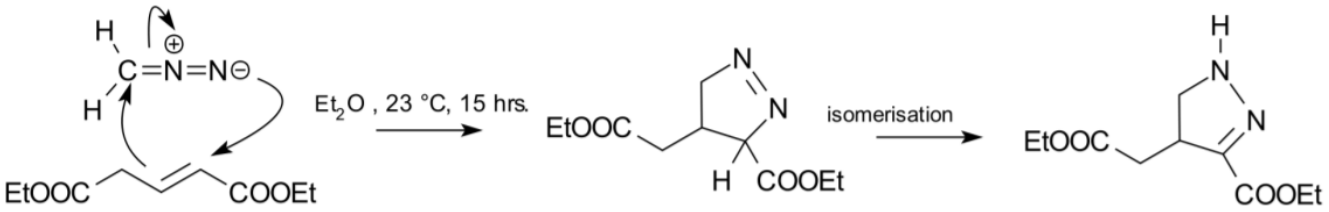


Figure 2.4.6.2.10 *A representation 1,3-dipolar cycloaddition to yield a simple diazoalkane is shown above.*
http://www.wikiwand.com/en/1,3-Dipolar_cycloaddition

Cope Rearrangement (RAR)

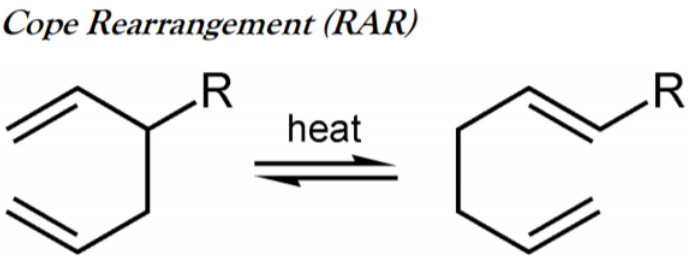


Figure 2.4.6.2.11 *The Cope depicted for 3-alkyl-1,5-hexadiene.*
http://www.wikiwand.com/en/Cope_rearrangement

The Cope RAR is an extensively studied [3,3]-sigmatropic RAR of 1,5-dienes developed by Arthur Cope. For example, 3-methyl-1,5-hexadiene heated to 300 °C yields 1,5- heptadiene (see Figure 2.4.6.2.11).

Oxidation/Reduction Reactions

The term “redox”, short for **red**uction–**ox**idation reaction, is used to describe the complete set of chemical reactions occurring for electron transfers leading to changes in oxidation states. Redox reactions require a reduction and complementary oxidation processes. The chemical species (reductant) where the electron is removed is oxidized, while the counter species (oxidant) to where it is added is reduced:

- Oxidation is the loss of electrons or increase in oxidation state;
- Reduction is the gain of electrons or a decrease in oxidation state.

As an example, in an electrolytic cell (shown in Figure 2.4.6.2.12), a complete redox (cycle) is shown.

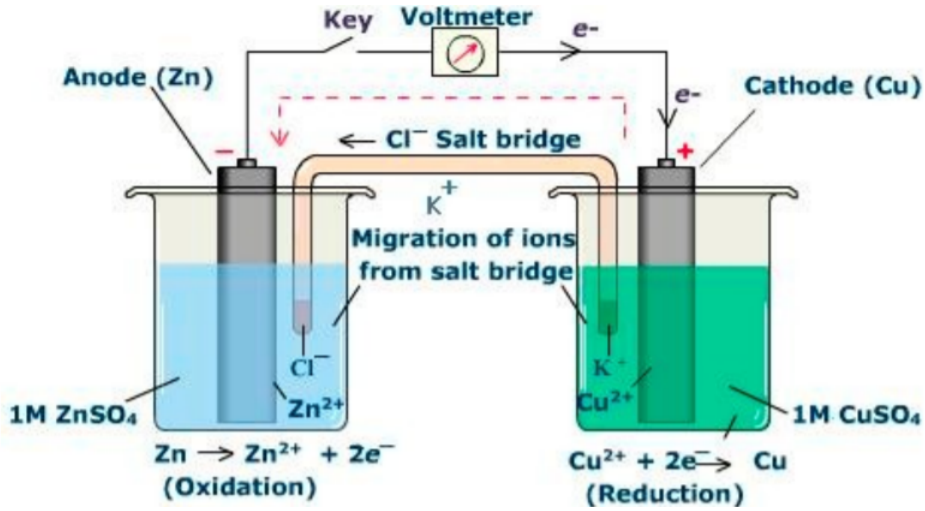


Figure 2.4.6.2.12 *A representation of the electron flow occurring during a redox cycle in which zinc is the reductant and copper the oxidant. commons.wikimedia.org/wiki/F...l_example.jpeg*

In the example, zinc is being oxidized from $Zn^0 \rightarrow Zn^{+2}$ which is a net loss of two electrons per atom. The oxidized zinc solubilizes in the anodic solution and proceeds to pull over an anion from the cathodic solution through the electromotive force of the two electrons that it lost. In other words, the zinc cation is balanced by the anions (likely two chloride anions) in the cathodic solution. Simultaneously, copper sulfate cations (cupric cations) are reduced by the net influx of the two electrons. A redox reaction can occur relatively slowly, for example, in the case of rust, or more rapidly, as in fires.

This page titled 2.4.6.2: Reaction Concepts is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **6.2: Reaction Concepts** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

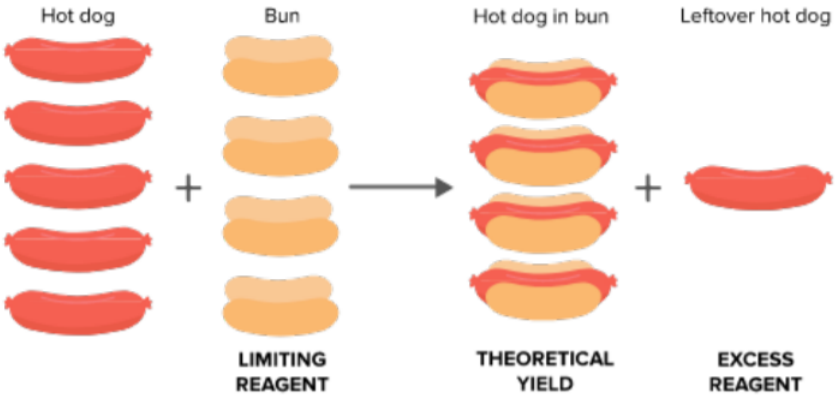
2.4.6.3: Reaction Design Concepts

Reaction Design and Efficiency

Green chemistry is all about increasing the overall efficiency of a reaction in terms of numerous criteria. In general, we try to mimic biological reactions because they are among the most efficient, low energy, and atom conserving reactions at our disposal. In terms of the "perfect" chemical reaction, we are talking about:

- *Selectivity:* we want to make one thing and one thing only to ensure that we don't have contamination or other issues to content with the require cleanup, purification, or multi- step pathways.
- *Efficiency*: This term refers to the overall energy, time, and material inputs relative to the final product required. How intense is the process needed to obtain a final product? Obviously, a more efficient process/reaction will require much less inputs. In an ideal scenario, a product occurs spontaneously without very little input.
- *Safety*: Although we don't talk about it much, the criterion of safety for a reaction or process must be paramount and very redundant given "Murphy's Law". Green chemistry is truly a discipline that is framed in the overall theme of safe performance and operation for a process or else it insists upon a no-go juncture point.
- *Solventless*: Again, the overall inputs are minimized in any reaction, and this is just one of the many criteria to ensuring efficiency and lowest impact. To do a reaction solventless means allowing self-reaction to occur with or a bimolecular reaction with other reagents in the absence of a solvent. Although many biological reactions occur in a solvent, ideally, it would be best to avoid if possible because of the many issues surrounding solvents: disposal, clean up, concentration-based reactions (depend on kinetics), and in the case of water, its precious nature for food, drinking, washing, etc.
- *Low or no energetics*: Here we try to approach an ideality for energy conservation in which we reduce the need for high heat, high pressure, long times, etc. Ideally, we would use the energy available to us without attempting to burn gas, coal, employ electricity (power grids), etc.
- *Chemical yield*: In most standard chemistry reaction descriptions, we tend to emphasize the "yield" of a reaction as the chief parameter for assessing the reaction efficiency or utility. A low yield reaction would be undesirable because of the lack of return on the process. Chemical yield is defined as the moles of desired products obtained as a ratio to the possible moles (theoretical) possible expressed as a percentage.

Let's give an excellent intuitive example from the Khan Academy (www.khanacademy.org/science/...-percent-yield): Let's assume you have five hot dogs and four hot dog buns. Obviously, you will have one hot dog in excess.



In terms of chemical reactions, the hot dog buns are the limiting reagent and the leftover single hot dog is the excess reagent; thus, four complete hot dogs are the theoretical yield, assuming the hot dogs and buns combine in a one-to-one ratio (you've got to balance the reaction first). In any chemical reaction, the limiting reagent is the reactant that determines how much of the products can be made. The other reactant(s) is/are in *excess*.

How do we measure yield more efficiently?

Typically, chemical yield focuses on a part of the reaction, i.e., our desired product based on the complete reaction of the limiting reagent. However, in many cases, the overall efficiency from a green chemical perspective is compromised despite high chemical

yields. Consider the "Gabriel Synthesis":

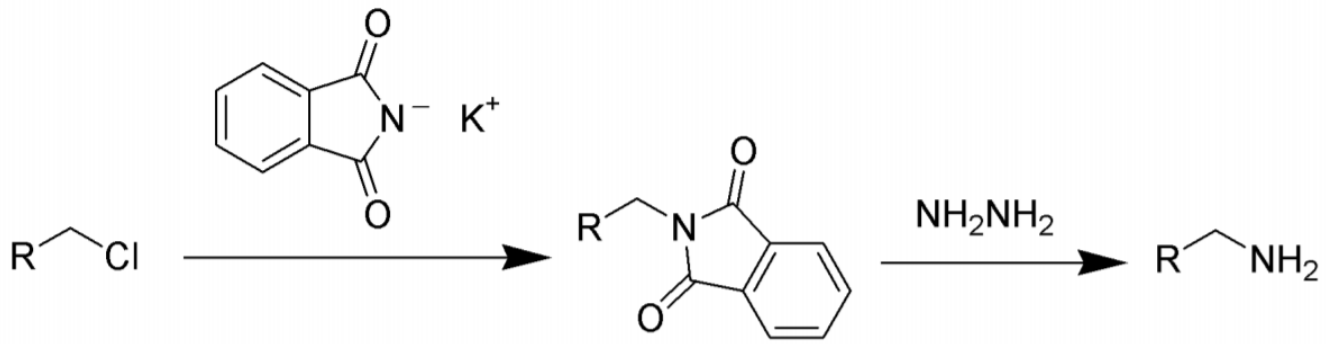


Figure 2.4.6.3.13: A representation of part of the classical Gabriel Synthesis whose product is a primary amine. http://www.wikiwand.com/en/Gabriel_synthesis

The reaction transforms primary alkyl halides into primary amines. In Figure 2.4.6.3.13, only the latter half of the reaction/synthesis is shown! Traditionally, the reaction uses potassium phthalimide (an $–NH_2$ synthon) to homologate the alkyl which is later coupled to the amine. In the Gabriel Synthesis, the phthalimide anion is employed as a surrogate of H_2N^-. The entirety of the reaction is shown in Figure 2.4.6.3.14 :

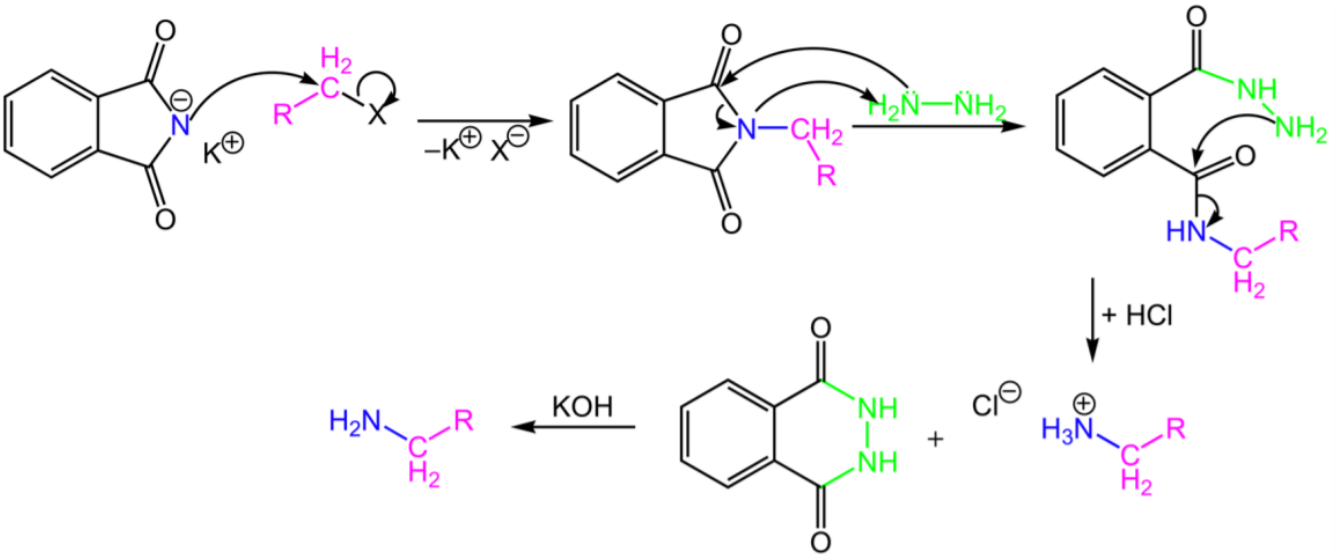


Figure 2.4.6.3.14: The Gabriel Synthesis that starts with the amine synthon (the first reactant at the left) that after four steps yields the desired primary amine. commons.wikimedia.org/wiki/File:Gabriel_synthesis_mechanism.png

In chemical yield methodologies, we do NOT account for the stream of by-products that amass during a reaction or set of reactions. In the above case, we have the following waste products: potassium halide (K-X), potassium chloride (KCl), and phthalhydrazide (see Figure 2.4.6.3.13 for the structure – the aromatic compound, second from right). Unbelievably, this array (especially the hydrazide) make up a huge fraction of wasted atoms that do not contribute in any way to the efficiency of the reaction. Because of the incredible "atomic" waste, we have a reaction with a very poor "atom economy".

Atom Economy

This is a concept that was first formalized by Trost. 1 Atom economy (atom efficiency) is the conversion efficiency of a chemical process in terms of all atoms and the desired product(s). It is one of the most singularly powerful definitions for understanding the "greenness" of a reaction and its potential usefulness.

Atom economy can be written as:

$$Atom\ Economy\ (\%) = \frac{Molecular\ Mass(MW)\ of\ Desired\ Product}{Molecular\ Mass(MW)\ of\ All\ Reactants} \times 100$$

As we already alluded, in an ideal chemical process, the amount of starting materials equals the amount of all products generated and no atom is wasted. If this is not the case, this is a genuine concern for raw materials that have a high cost or due to economic and environmental costs of disposal of the waste.

In addition to high-yielding process such as the Gabriel Synthesis that result in substantial byproducts, we also have the Cannizzaro Reaction[2] where 50% of the reactant aldehyde becomes the other oxidation state of the target and the Wittig[3], where high-mass phosphorus reagents are used but become waste.

A Diels-Alder reaction is an example of a very atom efficient reaction that also can be chemo-, regio-, diastereo- and enantioselective. Catalytic hydrogenation comes closest to being an ideal reaction that is extensively practiced both industrially and academically.[4]Poor atom economy is common in drug (pharmaceutical) synthesis, and in research. For example, during the synthesis of an alcohol by reduction of an ester with lithium aluminum hydride, the reaction produces a huge amount of aluminum salts. The cost can be considerable.

Experimental atom economy

In this derivative calculation of atom economy, we account for the actual quantities of reagents used in a ratio of theoretical yield of product total mass of reactants used as a percentage. Or equivalently, we can calculate the mass of reactants utilized in the desired product divided by the total mass of all reactants as a percentage

Percentage yield x experimental atom economy

This is considered the ultimate measure of the efficiency of a reaction. In this case, you multiply the chemical yield of the product by the experimental atom economy to obtain a true measure of the atom efficiency of the reaction

---

This page titled 2.4.6.3: Reaction Design Concepts is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

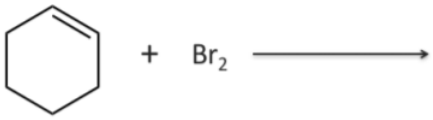
- **6.3: Reaction Design Concepts** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

2.4.6.4: Conclusions and Review

In this final chapter, we became very familiar with the concepts of reaction efficiency and design from a green chemistry perspective. We more or less employed the totality of the green chemistry principles to understanding what goes into maximizing the environmental, social, and economic benefits associated with scientific enterprises. We thoroughly explored the nature of the various types of reactions done in both fundamental and applied settings and how to evaluate them using very useful mathematical constructs. This chapter serves to cement the concepts we learned in this magnificent discipline and to provide a platform for understanding how to best approach the chemistry, chemical engineering, and the biological transformations of reactions and processes.

Review Questions

1. Indicate by drawings what the addition products are for the following reaction and what type of reaction it is:

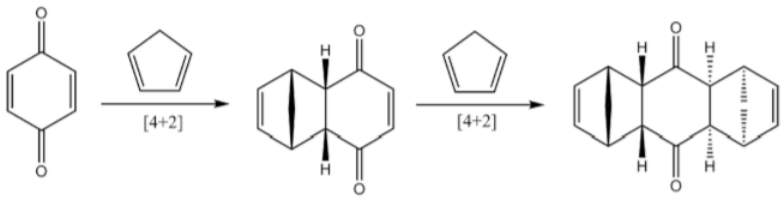


2. Please indicate what the products are of the reaction of 2-phenyl isopropyl iodide with sodium methoxide.

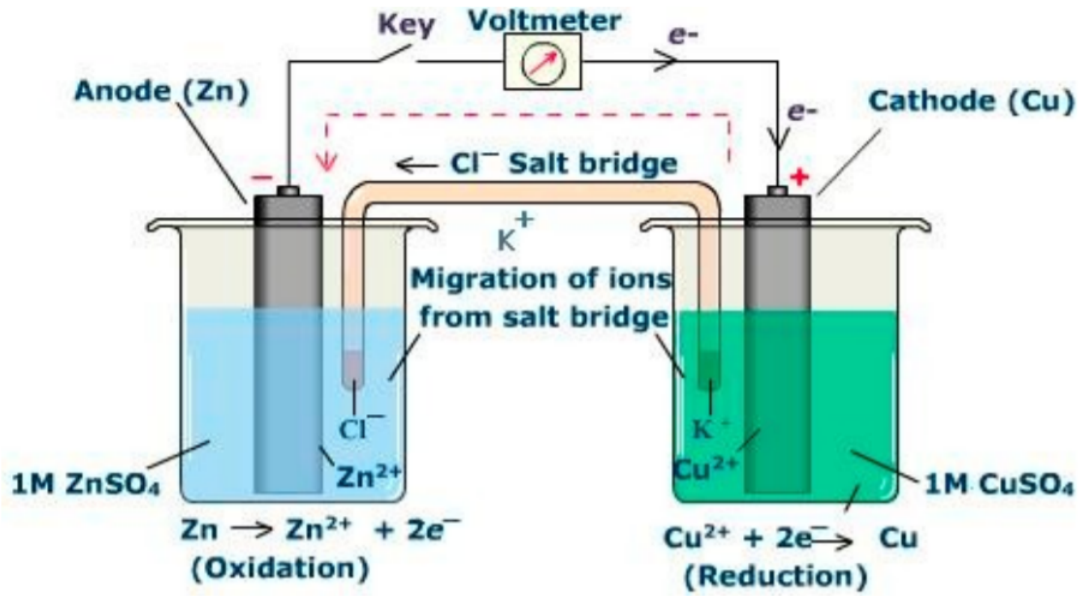
3. Does the non-salt product in (2) retain its stereochemical configuration? Why?

4. Please explain how the minor product is formed in the E2 reaction below and why it is minor?

5. Using the Diels-Alder Reaction shown below, draw the transition state that will lead to the adduct.



6. Using the redox diagram below, indicate why the process cannot go on forever



7. Define the following terms

a. atom economy

b. percent atom economy

c. percent yield

d. percent experimental atom economy

e. percent atom economy x experimental atom economy

8. Which of the terms in question 7 are only meaningful with experimental results? What experimental result is necessary to make these terms meaningful?

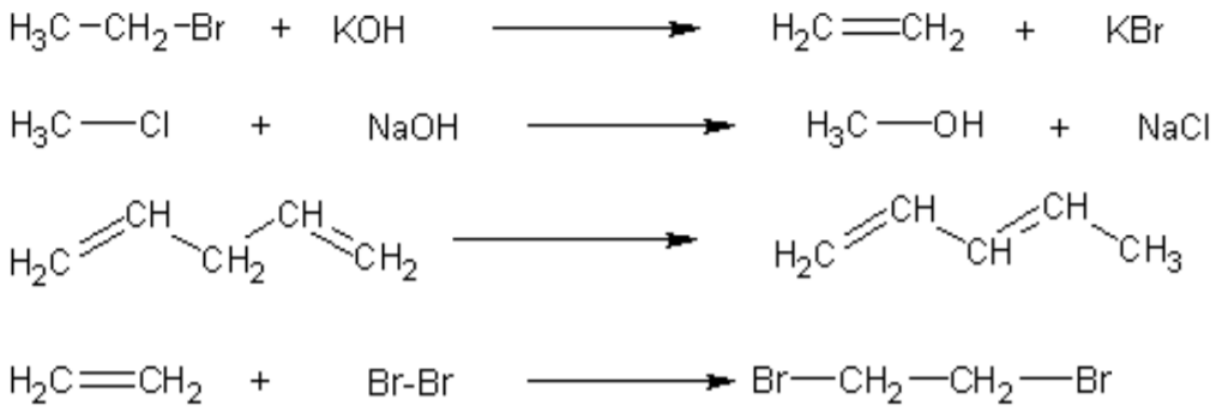
9. Consider the following four (4) reactions shown below:

a. Label each reaction as a substitution, elimination, addition or rearrangement.

b. Rewrite each reaction making sure that the reaction is balanced. Show all the reactant atoms that are incorporated into the desired product, and the atoms of the desired product in green, and all other atoms in brown.

c. Set up a table of atom economy whose headings are: formula, formula weight, number of utilized atoms, weight of utilized atoms, unutilized atoms, and weight of unutilized atoms.

d. Calculate the %-atom economy of each reaction



10. In the Further Reading section (#3), Roger Sheldon has developed a term very similar to the %-atom economy termed % atom utilization. The % atom utilization can be calculated according to the following equation: % Atom Utilization = (MW of desired product/MW of all products) X 100

a. Compare and contrast this with the %-atom economy

b. What concept that you learned in freshman chemistry makes the actual percentages calculated for the % atom utilization, and % atom economy equal (in most circumstances). Prove this by calculating the % atom utilization for each of the reactions in questions 9—13 and comparing your results to the %-atom economy that you calculated in question 9.

Further Reading

1. Williamson, Kenneth L., Macroscale and Microscale Organic Experiments, 2nd ed., D. C. Heath and Co., 1994, 247-252.
2. Trost, Barry M., The Atom Economy-A Search for Synthetic Efficiency. Science 1991, 254, 1471-1477.
3. Cann, Michael C.; Connelly, Marc E. Real-World Cases in Green Chemistry; ACS, Washington, 2000. Roger Sheldon of Delft University has developed a very similar concept called % atom utilization. Sheldon, Roger A. Organic Synthesis- Past, Present and Future. Chem. Ind. (London), 1992, (Dec), 903-906.
4. Anastas, Paul T., and Warner, John C. Green Chemistry Theory and Practice, Oxford University Press, New York, 1998.
5. Westheimer, T. The Award Goes to BHC. Chem. Eng. 1993, (Dec.), 84-95

This page titled 2.4.6.4: Conclusions and Review is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Lucian Lucia via source content that was edited to the style and standards of the LibreTexts platform.

- **6.4: Conclusions and Review** by Lucian Lucia is licensed CC BY 4.0. Original source: https://open.umn.edu/opentextbooks/formats/802.

CHAPTER OVERVIEW

Back Matter

Detailed Licensing

Detailed Licensing

TABLE OF CONTENTS

Green Chemistry seeks to present a body of chemical knowledge from the most fundamental level within a framework of the relationship of chemical science to human beings, their surroundings, and their environment. This book discusses real-world chemistry, introducing chemical principles as needed.

Topic hierarchy

Licensing

2.5.1: Sustainability and the Environment

- 2.5.1.1: Sustainability
- 2.5.1.2: The Environment and the Five Environmental Sphere
- 2.5.1.3: Seeing Green
- 2.5.1.4: Natural Capital of the Earth
- 2.5.1.5: Sustainability as a Group Effort- It Takes a (Very Big) Village
- 2.5.1.6: Sustainable Energy- Away from the Sun and Back Again
- 2.5.1.7: Green Science
- 2.5.1.8: Green Technology
- 2.5.1.9: Sustainability and the Eco-Economy
- 2.5.1.10: Green Products and Services- Design for Sustainability
- 2.5.1.11: Questions and Problems

2.5.2: The Key Role of Chemistry and Making Chemistry Green

- 2.5.2.1: Chemistry is Good (and Unavoidable)
- 2.5.2.2: The Environment and its Chemistry
- 2.5.2.3: What is Environmental Chemistry?
- 2.5.2.4: Environmental Pollution
- 2.5.2.5: Practice of Green Chemistry
- 2.5.2.6: Green Chemistry and Synthetic Chemistry
- 2.5.2.7: Reduction of Risk- Hazard and Exposure
- 2.5.2.8: The Risks of No Risks
- 2.5.2.9: Waste Prevention
- 2.5.2.10: Basic Principles of Green Chemistry
- 2.5.2.11: Some Things to Know About Chemistry before You Even Start
- 2.5.2.12: Combining Atoms to Make Molecules and Compounds
- 2.5.2.13: The process of making and breaking chemical bonds- chemical reactions
- 2.5.2.14: The Nature of Matter and States of Matter
- 2.5.2.15: References and Questions

2.5.3: The Elements - Basic Building Blocks of Green Chemicals

- 2.5.3.1: Elements, Atoms, and Atomic Theory
- 2.5.3.2: Hydrogen - The Simplest Atom
- 2.5.3.3: Helium - The First Noble Gas

2.5.3.4: Lithium, The First Metal
2.5.3.5: The Second Period of the Periodic Table
2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons
2.5.3.7: Completing the 20-Element Periodic Table
2.5.3.8: The Brief Periodic Table is Complete
2.5.3.E: The Elements - Basic Building Blocks of Green Chemicals (Exercises)
2.5.3.R: References

2.5.4: Compounds- Safer Materials for a Safer World

2.5.4.1: Chemical Bonds and Compound Formation
2.5.4.2: Electrons Involved in Chemical Bonds and Octets of Electrons
2.5.4.3: Sodium Chloride and Ionic Bonds
2.5.4.4: Covalent Bonds in H2 and Other Molecules
2.5.4.5: Covalent Bonds in Compounds
2.5.4.6: Covalent Bonds and Green Chemistry
2.5.4.7: Predicting Formulas of Covalently Bound Compounds
2.5.4.8: Chemical Formulas, the Mole, and Percentage Composition
2.5.4.9: What Are Chemical Compounds Called?
2.5.4.10: Acids, Bases and Salts
2.5.4.11: Questions and Problems

2.5.5: Chemical Reactions- Making Materials Safely and Sustainable

2.5.5.1: Describing What Happens with Chemical Equations
2.5.5.2: Balancing Chemical Equations
2.5.5.3: Just Because You Can Write It Doesn't Mean That It Will Happen
2.5.5.4: Yield And Atom Economy in Chemical Reactions
2.5.5.5: Catalysts That Make Reactions Go
2.5.5.6: Kinds of Chemical Reactions
2.5.5.7: Oxidation-Reduction Reactions and Green Chemistry
2.5.5.8: Quantitative Information from Chemical Reactions
2.5.5.9: Energy in Chemical Reactions
2.5.5.10: Stoichiometry by the Mole Ratio Method
2.5.5.11: Limiting Reactant and Percent Yield
2.5.5.12: Titrations - Measuring Moles by Volume of Solution
2.5.5.13: Industrial Chemical Reactions - The Solvay Process
2.5.5.14: Questions and Problems

2.5.6: The Wonderful World of Carbon - Organic Chemistry and Biochemicals

2.5.6.1: Rings and Chains of Carbon Atoms
2.5.6.2: Compounds of Carbon and Hydrogen - Hydrocarbons
2.5.6.3: Using Lines To Show Organic Structural Formulas
2.5.6.4: Functional Groups
2.5.6.5: Giant Molecules from Small Organic Molecules
2.5.6.6: Questions and Problems

2.5.6.7: Supplementary References

2.5.7: Chemistry of Life and Green Chemistry

2.5.7.1: Green Biochemistry
2.5.7.2: Biochemistry and the Cells
2.5.7.3: Carbohydrates
2.5.7.4: Proteins
2.5.7.5: Lipids - Fats, Oils, and Hormones
2.5.7.6: Nucleic Acids
2.5.7.7: Enzymes
2.5.7.8: Biochemical Processes in Metabolism
2.5.7.9: Biochemistry of Toxic Substances and Toxicological Chemistry
2.5.7.10: Literature Cited
2.5.7.11: Questions and Problems
2.5.7.12: Supplementary References

2.5.8: The Five Environmental Spheres and Biogeochemical Cycle

2.5.8.1: Five Environmental Spheres
2.5.8.2: The Hydrosphere
2.5.8.3: The Atmosphere
2.5.8.4: The Geosphere
2.5.8.5: The Biosphere
2.5.8.6: The Anthrosphere
2.5.8.7: Cycles of Matter
2.5.8.8: Questions and Problems
2.5.8.9: Supplementary References

2.5.9: Water - the Ultimate Green Substance

2.5.9.1: H_2O - Simple Formula, Remarkable Molecule
2.5.9.2: Occurrence, Availability , and Utilization of Water on Earth
2.5.9.3: Water Chemistry
2.5.9.4: Water Pollution
2.5.9.5: Greening of Water - Purification Before Use
2.5.9.6: Wastewater Treatment
2.5.9.7: Advanced Wastewater Treatment and Recycling
2.5.9.8: The Many Uses of Water
2.5.9.9: Hot Water - Pressurized Subcritical Water
2.5.9.10: Supercritical Water
2.5.9.11: Literature Cited
2.5.9.12: Questions and Problems
2.5.9.13: Supplementary References

2.5.10: Blue Skies for a Green Environment

2.5.10.1: More Than Just Air to Breathe
2.5.10.2: The Gas Laws

2.5.10.3: The Protective Atmosphere
2.5.10.4: Atmospheric Chemistry and Photochemical Reactions
2.5.10.5: The Enormous Importance of Climate
2.5.10.6: Atmospheric Particle Pollutants
2.5.10.7: Pollutant Gaseous Oxides
2.5.10.8: Acid Rain
2.5.10.9: Miscellaneous Gases in the Atmosphere
2.5.10.10: Photochemical Smog
2.5.10.11: Natural Capital of the Atmosphere
2.5.10.12: Energy and Mass Transfer in the Atmosphere
2.5.10.13: Questions and Problems
2.5.10.14: Supplementary References

2.5.11: The Geosphere and a Green Earth

2.5.11.1: Probing Deep into the Geosphere - A Well Too Far?
2.5.11.2: The Nature of the Geosphere
2.5.11.3: The Geosphere as a Source of Natural Capital
2.5.11.4: Environmental Hazards of the Geosphere
2.5.11.5: Water on and in the Geosphere
2.5.11.6: Anthrospheric Influences on the Geosphere
2.5.11.7: The Geosphere as a Waste Repository
2.5.11.8: Have You Thanked a Clod Today?
2.5.11.9: Plant Nutrients and Fertilizers in Soil
2.5.11.10: Soil and Plants Related To Wastes and Pollutants
2.5.11.11: Soil Loss - Desertification and Deforestation
2.5.11.12: Literature Cited and Supplementary References
2.5.11.13: Questions and Problems

2.5.12: The Biosphere and the Role of Green Chemistry in Feeding a Hungry World

2.5.12.1: Pigweed’s Revenge
2.5.12.2: The Biosphere
2.5.12.3: Cells - Basic Units of Life
2.5.12.4: Metabolism and Control in Organisms
2.5.12.5: Reproduction and Inherited Traits
2.5.12.6: Stability and Equilibrium of the Biosphere
2.5.12.7: DNA and the Human Genome
2.5.12.8: Genetic Engineering
2.5.12.9: Biological Interaction with Environmental Chemicals
2.5.12.10: Biodegradation
2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture
2.5.12.12: Agricultural Applications of Genetically Modified Organisms
2.5.12.13: The Anthrosphere in Support of the Biosphere
2.5.12.14: Livestock and their Wastes
2.5.12.15: Questions and Problems

2.5.12.16: Supplementary References

2.5.13: The Anthrosphere, Industrial Ecology, and Green Chemistry

2.5.13.1: Industrial Ecology and Industrial Ecosystems
2.5.13.2: Metabolic Processes in Industrial Ecosystems
2.5.13.3: Life Cycles in Industrial Ecosystems
2.5.13.4: Kinds of Products
2.5.13.5: Attributes Required by an Industrial Ecosystem
2.5.13.6: The Kalundborg Industrial Ecosystem
2.5.13.7: Environmental Impacts of Industrial Ecosystems
2.5.13.8: Green Chemistry and Industrial Ecology
2.5.13.9: Predicting and Reducing Hazards
2.5.13.10: The E-Factor in Green Chemistry
2.5.13.11: Catalysts and Catalysis
2.5.13.12: Biocatalysis with Enzymes
2.5.13.13: Energizing Chemical Reactions and Process Intensification
2.5.13.14: Solvents and Alternate Reaction Media
2.5.13.15: Feedstocks and Reagents
2.5.13.16: Literature Cited and Supplementary References
2.5.13.17: Questions and Problems

2.5.14: Feeding the Anthrosphere- Utilizing Renewable and Biological Materials

2.5.14.1: Feeding the Anthrosphere
2.5.14.2: Utilization of Feedstocks
2.5.14.3: Biological Feedstocks
2.5.14.4: Biological Sources of Chemicals
2.5.14.5: Biorefineries and Biomass Utilization
2.5.14.6: Monosaccharide Feedstocks - Glucose, Fructose, and Xylose
2.5.14.7: Cellulose
2.5.14.8: Feedstocks from Cellulose Wastes
2.5.14.9: Lignin
2.5.14.10: Direct Biosynthesis of Polymers
2.5.14.11: Bioconversion Processes for Synthetic Chemicals
2.5.14.12: Bamboo - Ancient Material for the Future
2.5.14.13: Literature Cited
2.5.14.14: Questions and Problems

2.5.15: Sustainable Energy- The Essential Basis of Green Systems

2.5.15.1: Sustainability and Energy
2.5.15.2: What is Energy?
2.5.15.3: Radiant Energy from the Sun
2.5.15.4: Sources of Energy Used in the Anthrosphere
2.5.15.5: Conversions Between Forms of Energy
2.5.15.6: Green Technology for Energy Conversion
2.5.15.7: Energy Conservation

2.5.15.8: Depletable Fossil Fuels
2.5.15.9: Carbon Sequestration for Fossil Fuel Utilization
2.5.15.10: Nuclear Energy
2.5.15.11: Renewable Energy Sources - Solar Energy
2.5.15.12: Energy from Wind and Water
2.5.15.13: Biomass Energy
2.5.15.14: Geothermal Energy
2.5.15.15: Hydrogen for Energy Storage and Utilization
2.5.15.16: Literature Cited and Supplementary References
2.5.15.17: Questions and Problems

2.5.16: Terrorism, Toxicity, and Vulnerability- Green Chemistry and Technology in Defense of Human Welfare

2.5.16.1: Vulnerability to Terrorist Attack
2.5.16.2: Protecting the Anthrosphere
2.5.16.3: Substances That Explode, Burn, or React Violently
2.5.16.4: Toxic Substances and Toxicology
2.5.16.5: Toxic Chemical Attack
2.5.16.6: Protecting Water, Food, and Air
2.5.16.7: Detecting Hazards
2.5.16.8: Green Chemistry to Combat Terrorism
2.5.16.9: Green Chemistry for Sustainable Prosperity and a Safer World
2.5.16.10: Green Chemistry to Combat Terrorism
2.5.16.11: Green Chemistry for Sustainable Prosperity and and a Safer World
2.5.16.12: Literature Cited
2.5.16.13: Questions and Problems

2.5.17: The Ten Commandments of Sustainability and Sensible Measure

2.5.17.1: We Cannot Go on Like This
2.5.17.2: The First Commandment
2.5.17.3: The Second Commandment
2.5.17.4: The Third Commandment
2.5.17.5: The Fourth Commandment
2.5.17.6: The Fifth Commandment
2.5.17.7: The Sixth Commandment
2.5.17.8: The Seventh Commandment
2.5.17.9: The Ninth Commandment
2.5.17.10: The Tenth Commandment
2.5.17.11: Some Sensible Measures for Sustainability
2.5.17.12: The Eighth Commandment
2.5.17.13: Literature Cited
2.5.17.14: Questions and Problems
2.5.17.15: Supplementary References

Detailed Licensing

This page titled 2.5: Green Chemistry and the Ten Commandments of Sustainability (Manahan) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

CHAPTER OVERVIEW

Front Matter

Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

SECTION OVERVIEW

2.5.1: Sustainability and the Environment

2.5.1.1: Sustainability

2.5.1.2: The Environment and the Five Environmental Sphere

2.5.1.3: Seeing Green

2.5.1.4: Natural Capital of the Earth

2.5.1.5: Sustainability as a Group Effort- It Takes a (Very Big) Village

2.5.1.6: Sustainable Energy- Away from the Sun and Back Again

2.5.1.7: Green Science

2.5.1.8: Green Technology

2.5.1.9: Sustainability and the Eco-Economy

2.5.1.10: Green Products and Services- Design for Sustainability

2.5.1.11: Questions and Problems

"If we do not change direction, we are likely to end up where we are headed," (old Chinese proverb)."If we make the effort to learn its language, Earth will speak to us and tell us what we must do to survive."

This page titled 2.5.1: Sustainability and the Environment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.1.1: Sustainability

The old Chinese proverb certainly applies to modern civilization and its relationship to world resources that support it. Evidence abounds that humans are degrading the Earth life support system upon which they depend for their existence. The emission to the atmosphere of carbon dioxide and other greenhouse gases is almost certainly causing global warming and climate change. Discharge of pollutants has degraded the atmosphere, the hydrosphere, and the geosphere in industrialized areas and has placed great stress on parts of the biosphere. Natural resources including minerals, fossil fuels, fresh water, and biomass have become stressed and depleted. The productivity of agricultural land has been diminished by water and soil erosion, deforestation, desertification, contamination, and conversion to non-agricultural uses. Wildlife habitats including woodlands, grasslands, estuaries, and wetlands have been destroyed or damaged. About 3 billion people (half of the world's population) live in dire poverty on less than the equivalent of U.S. $2/day. The majority of these people lack access to sanitary sewers and the conditions under which they live give rise to debilitating viral, bacterial, and protozoal diseases. At the other end of the living standard scale, a relatively small fraction of the world's population consumes an inordinate amount of resources with lifestyles that involve living too far from where they work in energy-wasting houses that are far larger than they need, commuting long distances in large "sport utility vehicles" that consume far too much fuel, and overeating to the point of unhealthy obesity with accompanying problems of heart disease, diabetes, and other obesity-related maladies.

As We Enter the Anthropocene

Humans have gained an enormous capacity to alter Earth and its support systems. Their influence is so great that we are now entering a new epoch, the **Anthropocene**, in which human activities have effects that largely determine conditions on the planet. The major effects of humans upon Earth have taken place within a minuscule period of time relative to that during which life has been present on the planet or, indeed, relative to the time that modern humans have existed. These effects are largely unpredictable, but it is essential for humans to be aware of the enormous power in their hands — and of their limitations if they get it wrong and ruin Earth and its climate as life support systems.

Achieving Sustainability

Although the condition of the world and its human stewards outlined above sounds rather grim and pessimistic, this is not a grim and pessimistic book. That is because the will and ingenuity of humans that have given rise to conditions leading to deterioration of Planet Earth can be — indeed, are being — harnessed to preserve the planet, its resources, and its characteristics that are conducive to healthy and productive human life. The key is **sustainability** or **sustainable development** defined by the Bruntland Commission in 1987 as *industrial progress that meets the needs of the present without compromising the ability of future generations to meet their own needs*. 1A key aspect of sustainability is the maintenance of Earth's **carrying capacity**, that is, its ability to maintain an acceptable level of human activity and consumption over a sustained period of time. Although change is a normal characteristic of nature, sudden and dramatic change can cause devastating damage to Earth support systems. Change that occurs faster than such systems can adjust can cause irreversible damage to them. In addition to its main theme of green chemistry, a major purpose of this book is to serve as an overview of the science and technology of sustainability emphasizing sustainable chemistry as well as the general science and technology of sustainability.

Rethinking Environmentalism and Sustainability

The common view of a good, sustainable environment as a rural, low-population-density area may be misleading. A convincing argument for this proposition is made in the 2009 book *Green Metropolis: Why Living Smaller, Living Closer, and Driving Less are the Keys to Sustainability*.2Classified as an "eco-urbanist manifesto," this book makes the somewhat surprising case that New York City's Manhattan is a model of sustainability for the modern overpopulated world. This densely populated compact city emits less than one third as much greenhouse gas per person compared to the average for the United States. One reason why this is so is that the large apartment buildings and other large structures in New York City are very efficient in conserving heat; that which leaks from one tends to end up heating another. Cold air produced by air conditioning in the summer is similarly conserved. Another reason that the city is energy-efficient stems from its outrageously congested traffic and lack of affordable parking meaning that the automobile is impractical for most residents thereby forcing reliance on far more efficient public transportation. Only about one-fifth of New York City's residents regularly commute with individual automobiles. In contrast, those who live "close to nature" in

rural settings tend to dwell in free-standing houses that are inherently less energy efficient than apartment buildings and by necessity they must commute with energy-wasting vehicles. If they live on unimproved roads they may require especially inefficient large, rugged four-wheel-drive vehicles. Compensation cannot be made for such a lifestyle by measures advocated by many environmentalists, such as backyard compost piles and fuel-efficient vehicles.

According to *Green Metropolis*, New York City, which has a population density more than 800 times that of the U.S. as a whole and about 30 times that of Los Angeles, offers a model for a growing world population to exist within the confines of Earth's limited resources. The prescription for sustainability is to "live smaller, live closer, and drive less." To that may be added "reproduce less" in that dense urban environments tend to discourage large families. A major culprit in the development of modern environmental problems is the public obsession with the private automobile, which enables destructive urban sprawl and excessive consumption of gasoline. One of the unintended consequences of the laudable goal of increased fuel economy in automobiles is to make them more affordable to use, thus facilitating destructive urban sprawl. The automobile-based societies of the U.S. and many other industrialized nations has been made possible by the exploitation of relatively abundant and inexpensive petroleum. In years to come, as petroleum inevitably becomes more scarce and expensive, these societies will have to undergo wrenching changes, the best end result of which would be much more sustainable, compact urban societies

---

This page titled 2.5.1.1: Sustainability is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **1.1: Sustainability** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.1.2: The Environment and the Five Environmental Sphere

Since this book deals with the environment, it is important to know what is meant by the environment. Essentially, the **environment** consists of our surroundings, which may affect us and which, in turn, we may affect. A part of the environment may consist of rock formations several kilometers below Earth's surface, so deep that humans cannot reach them and of which they are scarcely aware except in those instances when the rock formations shift along a fault line and cause an earthquake, which may be very destructive and take many lives. Another part of the environment is the atmosphere touching Earth's surface, a part of the environment with which humans are always in contact and which is essential for the life-giving oxygen that they require. In discussing the environment it is helpful to regard it as consisting of the following five spheres as shown in Figure 1.1: (1) The **hydrosphere**, (2) the **atmosphere**, (3) the **geosphere**, (4)the **biosphere**, and (5) the **anthrosphere** (that part of the environment constructed and operated by humans). These spheres overlap and interact strongly. For example, fish are part of the biosphere, dwell in the hydrosphere, and acquire dissolved oxygen that they need from the atmosphere. Mineral nutrients required by the fish and by the algae upon which the fish feed come from the geosphere. The part of the hydrosphere in which the fish reside may be a reservoir constructed by impounding a stream with a dam that is part of the anthrosphere. Many other such examples may be cited.

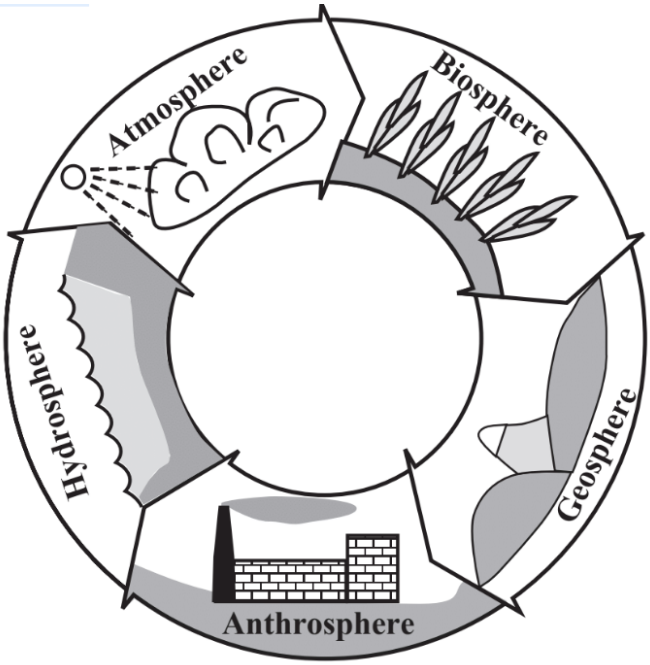


Figure 1.1. The environment may be considered as consisting of five spheres that exchange materials and energy, largely through biogeochemical cycles.

Biogeochemical cycles describe the exchange of materials among the five environmental spheres. Aspects of these environmentally crucial cycles are covered in various parts of this book. As the name implies, these cycles involve biological and geochemical phenomena but may also include processes that occur in the atmosphere and the hydrosphere as well as human influences on them. An important part of these cycles consists of the interfaces between the environmental spheres. The interfaces are often very thin with respect to Earth's whole environment. An important example of such an interface is the one between the geosphere and the atmosphere where most plants grow that support life on Earth. Typically this region extends into the geosphere for only the meter or less penetrated by plant roots and into the atmosphere only to the height of the plants. Within this region there are other interfaces including the biosphere/geosphere boundary between plant roots and soil and the biosphere/atmosphere boundary across which oxygen and carbon dioxide gas are exchanged between leaf surfaces and the atmosphere.

The study of the environment is **environmental science**, in its broadest sense the science of the complex interactions that occur among the terrestrial, atmospheric, aquatic, living, and anthropological systems that compose Earth and the surroundings that may affect living things.[3] It includes all the disciplines, such as chemistry, biology, ecology, sociology, and government, that affect or describe these interactions. For the purposes of this book, environmental science will be defined as the study of the earth, air, water, and living environments, and the effects of technology there on. To a significant degree, environmental science has evolved from investigations of the ways by which, and places in which, living organisms carry out their life cycles. This discipline used to be

known as natural history, which later evolved into ecology, the study of environmental factors that affect organisms and how organisms interact with these factors and with each other.

This page titled 2.5.1.2: The Environment and the Five Environmental Sphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **1.2: The Environment and the Five Environmental Sphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.1.3: Seeing Green

Given the dependence of humans upon a livable environment, it is essential that it be maintained in a healthy state. The maintenance of a healthy environment is commonly termed **sustainability**. In recent years there has been a lot of activity in the area of sustainability. Earlier efforts in the sustainability arena centered around pollution and its effects. Degradation of the environment has been a concern of thoughtful people for many decades. Dating back to the early 1800s and even before, the widespread use of high-sulfur coal for fuel was noted as a cause of bad air quality and impaired visibility in urban areas such as London. Water polluted by pathogenic microorganisms sickened and killed millions of people for many centuries. By the end of World War II, the atmosphere of Los Angeles had become noxious, irritating, and unhealthy due to the presence of ozone and other chemical oxidants, aldehydes, and small particulate matter. In some respects, this condition resembled pollution of the London atmosphere observed earlier, which was a combination of smoke and fog which some called “smog.” So the condition afflicting Los Angeles and a number of similar cities came to be known as smog, but a kind of smog that developed in air having low humidity and exposed to intense sunlight, conditions opposite of those under which London smog was formed. Chemically, the two kinds of smog were totally different in that London smog was in a reducing atmosphere with high concentrations of chemically reducing SO_2 whereas the Los Angeles smog is oxidizing and any SO_2 emitted to it is rapidly oxidized to sulfuric acid.

Concern over deterioration of the environment increased with the 1962 publication of Rachel Carson’s classic book *Silent Spring*,[4] the theme of which was that DDT and other mostly pesticidal chemicals were becoming concentrated through the food chain with the result that birds at the top of the chain, such as eagles and hawks, were producing eggs with soft shells that failed to produce viable baby birds. The implication was that substances harming bird populations might harm humans as well.

By about 1970 it was generally recognized that air, water, and land pollution was reaching intolerable levels. As a result, various countries passed and implemented laws designed to reduce pollutants and to clean up waste chemical sites at a cost that has easily exceeded one trillion dollars globally.

More recently concern over environmental pollution has extended beyond a narrow focus upon pollution and its effects to include the broader area of sustainability. The achievement of sustainability certainly requires avoiding pollution and counteracting its effects. But it also mean maintaining flows of essential materials, energy, food, safe water, healthy air, and the other things that humans and other organisms on Planet Earth require for their survival and well-being.

The term “green” has come to stand for sustainability in its various forms and is used throughout this book. Most of sustainability has to do with matter, and chemistry is the science of matter. It is only natural, therefore, that sustainable chemistry is now known as green chemistry, a discipline that has developed rapidly since about the mid-1990s. This book is about green chemistry. But the practice of green chemistry involves more broadly green science and technology, which are discussed in this book and related to green chemistry.

This page titled 2.5.1.3: Seeing Green is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **1.3: Seeing Green** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.1.4: Natural Capital of the Earth

All of that very small group of humans who have been privileged to view Earth from outer space have been struck with a sense of awe at the sight. Photographs of Earth taken at altitudes high enough to capture its entirety reveal a marvelous sphere, largely blue in color, white where covered by clouds, with desert regions showing up in shades of brown and red. But Earth is far more than a beautiful globe that inspires artists and poets. In a very practical sense it is a source of the life support systems that sustain humans and all other known forms of life. Earth obviously provides the substances required for life including water, atmospheric oxygen, carbon dioxide from which billions of tons of biomass are made each year by photosynthesis, and ranging all the way down to the trace levels of micronutrients such as iodine and chromium that organisms require for their metabolic processes. But more than materials are involved. Earth provides temperature conditions conducive to life and a shield against incoming ultraviolet radiation, its potentially deadly photons absorbed by molecules in the atmosphere, their energy dissipated as heat. Earth also has a good capacity to deal with waste products that are discharged to the atmosphere, into water, or into the geosphere.

The capacity of Earth to provide materials, protection, and conditions conducive to life is known as its **natural capital,** which can be regarded as the sum of two major components: **natural resources** and **ecosystem services**. These conditions are giving rise to a new business model termed **natural capitalism**. Early hunter-gatherer and agricultural human societies made few demands upon Earth's natural capital. As the industrial revolution developed from around 1800, natural resources were abundant and production of material goods was limited largely by labor and the capacity of machines to process materials. But now population is in excess, computerized machines have an enormous capacity to process materials, and the availability of natural capital is the limiting factor in production including availability of natural resources, the vital life-support ability of ecological systems, and the capacity of the natural environment to absorb the byproducts of industrial production, most notably greenhouse gas carbon dioxide.

Rather than the adversarial relationship that has prevailed between the traditional business community and environmentalists with regard to economic development, a functioning system of natural capitalism properly values natural and environmental resources. The goal of natural capitalism is to increase well-being, productivity, wealth, and capital while reducing waste, consumption of resources, and adverse environmental effects. The traditional capitalist economic system has proven powerful in delivering consumer goods and services using the leverage of individual and corporate incentives. A functional system of natural capitalism retains these economic drivers while incorporating sustainable practices such as recycling wastes back into the raw material stream and emphasizing the provision of services rather than just material goods. In so doing a system of natural capitalism emulates nature's systems through the practice of industrial ecology, discussed in detail in Chapter 14, and the application of the principles of green chemistry (see Chapter 2).

The development of a functional system of natural capitalism requires several important changes in business practices. These include the following:

1. Implement technologies that are highly productive with greatly reduced use of nonrenewable minerals and energy.
2. Develop systems in which waste materials and energy from one sector are utilized by another sector (a functioning system of industrial ecology).
3. Change business models from those that emphasize selling goods to those that concentrate upon providing services, for example, by selling fewer automobiles and providing better mass transportation systems.
4. Reinvest in natural capital to increase production of ecosystem services. An example is the provision of constructed wetlands as part of wastewater treatment systems to provide wildlife breeding grounds along with finishing of treated wastewater effluent.

Evolution of the Utilization of Natural Capital

Figure 2.5.1.4.1 shows the burden on Earth's natural capital as a function of the progression of economic development. During pre-industrial times the capacity of humans to deplete natural capital was minimal, largely because of limitations on the rates at which energy could be used. As the industrial revolution developed and humans learned how to harness energy sources, particularly from fossil fuels, Earth's natural capital was increasingly consumed in areas such as exploitation of depleting resources and utilization of the hydrosphere, geosphere, and atmosphere for the disposal of wastes. As industrialization progressed, it became increasingly obvious that it was causing problems in areas such as air and water pollution, soil erosion exacerbated by the capabilities of fossil-fueled tillage machinery to disturb soil, and depletion of rich ore sources necessitating the mining of much larger amounts of less rich ores to obtain needed quantities of

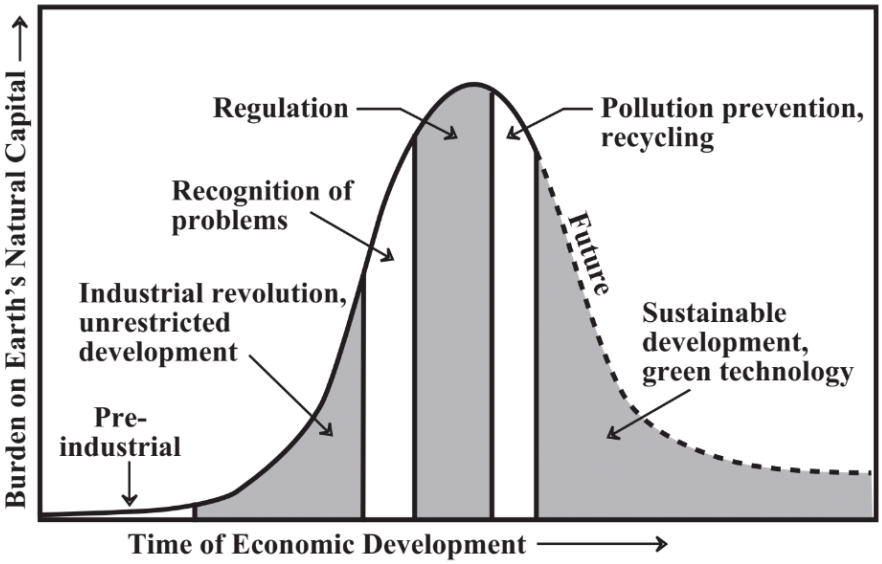


Figure 2.5.1.4.1. Stages of economic development with respect to utilization of Earth's natural capital.

metals and other geospheric resources. As a consequence, laws were passed and regulations put into place to reduce pollution and to conserve resources. Particularly after the devastating Dust Bowl years of the 1930s in which much productive topsoil was lost to wind erosion, the U.S.government initiated programs of soil conservation with incentives to preserve the essential soil resource. Efforts to reduce air and water pollution concentrated initially on the most obvious pollutants, such as particles emitted from smokestacks, followed by greater emphasis upon more insidious pollutants, such as heavy metals in water. The regulatory approach has been evolving into one that emphasizes pollution prevention, recycling, and conservation of energy and materials. A final phase is sustainable development and utilization of green technology that can support growing economic development while substantially reducing exploitation of Earth's natural capital.

This page titled 2.5.1.4: Natural Capital of the Earth is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **1.4: Natural Capital of the Earth** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.1.5: Sustainability as a Group Effort- It Takes a (Very Big) Village

The achievement of sustainability and preservation of natural capital requires intense efforts by both individuals and groups. This was illustrated centuries ago in England by "the tragedy of the commons"[5]. The commons consisted of a pasture shared by village residences to provide forage for their cows, sheep, and horses. An individual family could increase its wealth (in meat, milk, or horsepower) by adding an animal. For example, a one-cow family could double its wealth in cows by buying another and putting it to graze in the pasture. If the pasture was accommodating 100 cows, for example, this would have an apparent cost of only about 1% for the small community as a whole. The natural tendency was for families to keep on adding cows until a point was reached at which the pasture became exhausted and unproductive due to overgrazing, the animals died or had to be slaughtered, and the entire support system to provide milk and meat based upon the natural capital of the pasture in the commons collapsed. During the fourteenth century this unfortunate circumstance became so widespread that the economies of many villages collapsed with whole populations no longer able to provide for their basic food needs.

History has many examples of the tragedy of the commons. As an example, when the settlers began to cultivate what was formerly open rangeland in Edwards County, Texas, in the 1880s the ranchers who had used it for pasture met and proclaimed the following: "Resolved that none of us know, or care to know, anything about grasses, native or otherwise, outside of the fact that for the present, there are lots of them, the best on record, and we are getting the most of them while they last."[6] Soon the combined effects of overgrazing and drought reduced the yield of grass such that the ranchers' livelihoods were threatened and the newly cultivated land became unproductive. Shortsighted attitudes towards Earth's natural capital similar to those expressed by the ranchers continue to lead to many tragedies of the commons. In modern times heavy cultivation of marginal land is turning large areas to desert (desertification), the Amazon rain forest is being cut down and burned to provide a one-time harvest of wood and a few years of crop production (deforestation), severe deterioration of the global ocean fisheries resource is occurring, congested freeways at times become great linear parking lots and, of much direct concern to many university students and faculty, some parking facilities have become so oversold that their utility is seriously curtailed as paying customers cannot find parking space.

The logic of the commons holds true in modern times in which the global commons consist of the air humans must breathe, water resources, agricultural lands, mineral resources, capacity of the natural environment to absorb wastes, and all other facets of natural capital. According to the logic of the commons, each consumer has the right to acquire a segment of natural capital, the cost of which is distributed throughout the commons and shared by all. The natural competition among consumers results in some consumers acquiring relatively more of Earth's natural capital and becoming wealthier. Within limits this is a healthy consequence of capitalist systems. However, if enough consumer units use too much natural capital, it becomes exhausted and unsustainable, therefore unable to support the society as a whole, so that all suffer, including those on top of the consumer food chain.

Automotive transportation illustrates a modern tragedy of the commons. Acquisition of an automobile adds to an individual's possessions and mobility. The materials required to make a single automobile, the fuel to run it, and its exhaust pollutants make a minuscule impression on Earth's natural capital. However, when millions of people acquire automobiles, the demand on Earth's natural capital of materials, fuel, and ability to absorb pollutants becomes severely stressed, heavy traffic turns the automobile from a convenience into a burden, and, in some places at some times, the whole transportation system collapses.

These "tragedies of the commons" illustrate the limitations of unregulated "free-for-all" capitalist economic systems in achieving sustainable development and make a strong case for collective actions in the public sector to ensure that humankind can exist within the limits of Earth's natural capital. However, the collapse of Communist economic systems around 1990 left a legacy of abandoned, inefficient factories, poverty, and environmental degradation showing the adverse effects from discouraging private enterprise. In addition to enlightened regulations that ensure preservation of Earth's essential support systems, successful economic systems require human ingenuity, initiative, and even greed. Getting these and other incentives to work well on a planet in which natural capital is the major limiting economic factor is the huge challenge facing modern and developing economies.

There is an old African proverb that translates to, "It takes a village to raise a child." The idea is of course that successful child-rearing requires the efforts of more than just parents, but requires the efforts of an entire village. The same principle applies to Planet Earth except that in this case billions of children are being raised and it will take the efforts of a very large village — the population of the entire world — to preserve the planet and its resources upon which those billions of children must depend for their existence.

This page titled 2.5.1.5: Sustainability as a Group Effort- It Takes a (Very Big) Village is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **1.5: Sustainability as a Group Effort- It Takes a (Very Big) Village** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.1.6: Sustainable Energy- Away from the Sun and Back Again

As discussed in more detail in Chapter 15, the key to sustainability is abundant, environmentally safe energy. The evolution of humankind's utilization of energy is illustrated in Figure 2.5.1.6.1. Until very recently in the history of humankind we have depended upon the sun to meet our energy needs. The sun has kept most of the land mass of Earth at a temperature that enables human life to exist. Solar radiation has provided the energy for photosynthesis to convert atmospheric carbon dioxide to plant biomass providing humans with food, fiber, and wood employed for dwelling construction and fuel. Animals feeding on this biomass provided meat for food and hides and wool that humans used for clothing. Eventually humans developed means of indirectly using solar energy. This was especially true of wind driven by solar heating of air masses and used to propel sailing vessels and eventually to power windmills employed for power. The solar-powered hydrologic cycle provided flowing water, the energy of which was harnessed by water wheels. Virtually all the necessities of life came from utilization of solar energy.

The Brief Era of Fossil Fuels

Dating from around 1800, humankind began to exploit fossil fuels for their energy needs. Initially, coal was burned for heating and to power newly developed steam engines for mechanical energy used in manufacturing and steam locomotives. After about 1900 petroleum developed rapidly as a source of fuel and, with the development of the internal combustion engine, became the energy source of choice for transportation needs. Somewhat later natural gas developed as an energy source. The result was a massive shift from solar and biomass energy sources to fossil fuels. Utilization of fossil carbon-based materials resulted in a revolution that went far beyond energy utilization. One important example was the invention by Carl Bosch and Fritz Haber in Germany in the early 1900s, of a process for converting atmospheric elemental nitrogen from air to ammonia, NH_3 by the reaction of N_2 with H_2. This high-pressure, high-temperature process required large amounts of fossil fuel to provide energy and to react with steam to produce elemental hydrogen. The discovery of synthetic nitrogen fixation enabled the production of huge quantities of relatively inexpensive nitrogen fertilizer and the resulting increase in agricultural production may well have saved Europe, with a rapidly developing population at the time, from widespread starvation. (It also enabled the facile synthesis of great quantities of nitrogen-based explosives that killed millions of people in World War I and subsequent conflicts.) Fossil fuel, which has been described as "fossilized sunshine,"[7] resulted in an era of unprecedented material prosperity and an increase in human population from around 1 billion to over 6 billion.

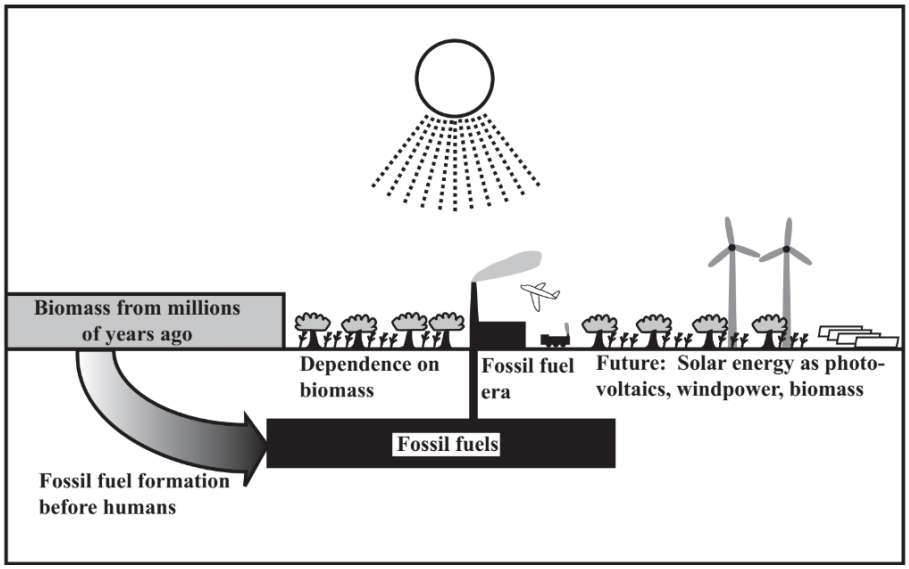


Figure 2.5.1.6.1, Evolution of the use of energy from solar and biomass sources through the brief, but spectacular, era of fossil fuels and on to renewable solar-based sources

By the year 2000 it had become obvious that the era of fossil fuels was not sustainable. One reason is that fossil fuel is a depleting resource that cannot last indefinitely as the major source of energy for the industrial society to which it has led. Approximately half of the world's total petroleum resource has already been consumed so that petroleum will continue to become more scarce and expensive and can last for only a few more decades as the dominant fuel and organic chemicals raw material. Coal is much more

abundant, but its utilization leads to the second reason that the era of fossil fuels must end because it is the major source of anthropogenic atmospheric carbon dioxide, greatly increased levels of which will almost certainly lead to global warming and massive climate change. Natural gas (methane, CH_4) is an ideal, clean-burning fossil fuel that produces the least amount of carbon dioxide per unit energy generated. Rapidly expanding new discoveries of natural gas largely from previously inaccessible tight shale formations means that it can serve as a "bridging fuel" for several decades until other sources can be developed. Nuclear energy, properly used with nuclear fuel reprocessing, can take on a greater share of energy production, especially for base load electricity generation. But clearly drastic shifts must occur in the ways in which energy is obtained and used.

Back to the Sun

With the closing of the brief but spectacular era of fossil hydrocarbons, the story of humankind and its relationship to Planet Earth is becoming one of "from the sun to fossil fuels and back again" as humankind returns to the sun as the dominant source of energy and photosynthetic energy to convert atmospheric carbon dioxide to biomass raw materials. In addition to direct uses for solar heating and for photovoltaic power generation, there is enormous potential to use the sun for the production of energy and material. Arguably the fastest-growing energy source in the world is wind-generated electricity. The wind is produced when the sun heats masses of air causing the air to expand. Once the dominant source of energy and materials, biomass produced by solar-powered photosynthesis is beginning to live up to its potential as a source of feedstocks to replace petroleum in petrochemicals manufacture and of energy in synthetic fuels (see Chapter 14,"Chapter 14 Feeding the Anthrosphere: Utilizing Renewable and Biological Materials" and Chapter 15, "Sustainable Energy: The Essential Basis of Sustainable Systems"). Biomass is still evolving as a practical source of liquid fuels. The two main ones of these are fermentation to produce ethanol and synthesis of biodiesel fuel made from plant lipid oils. Unfortunately, although ethanol made from sugar derived from sugar cane that grows prolifically in some areas such as Brazil is an economical gasoline substitute, the net energy gain from ethanol derived from cornstarch relies on the grain, the most valuable part of the plant otherwise used for food and animal feed; the net energy gain is marginal. The economics of producing synthetic biodiesel fuel from sources such as soybeans may be somewhat better. However, production of this fuel from oil palm trees in countries such as Malaysia is resulting in destruction of rain forests and diversion of palm oil from the food supply.

Practical means do exist to utilize biomass for energy and materials without seriously disrupting the food supply. Arguably the best way to do that is to thermochemically convert biomass to synthesis gas, a mixture of CO and H_2 that can be combined chemically by long-established synthetic routes to produce methane, larger-molecule hydrocarbons, alcohols and other products (see Chapter 15). The main pathway for so doing is to utilize biomass from renewable non-food biosources, which include crop byproducts (wheat straw, rice straw, cornstalks produced in surplus during the production of grain) and dedicated crops among which are highly productive hybrid poplar trees and sawgrass. Microscopic algae are especially promising as a biomass source because of their much higher productivity than terrestrial plants, their ability to grow in brackish(somewhat saline) water in containments in desert areas, and their ability to utilize sewage as a nutrient source. When biomass is used to produce synthesis gas, the essential nutrients, especially potassium and phosphorus, can be reclaimed from the biomass residues and used as fertilizer material to promote the growth of additional biomass.

Future scientific discoveries and technological advances will play key roles in the achievement of energy sustainability. Three areas in which Nobel-level breakthroughs are needed in the achievement of energy sustainability were expressed in a February 2009 interview by Dr.Steven Chu, a Nobel Prize winning physicist who had just been appointed Secretary of Energy inU.S. President Barack Obama's new administration. The first of these is in solar power in which the efficiency of solar energy capture and conversion to electricity needed to improve several-fold. A second area of need is for improved electric batteries to store electrical energy generated by renewable means and to enable practical driving ranges in electric vehicles. A third area in need of a quantum leap is for improved crops capable of converting solar energy to chemical energy in biomass by photosynthesis at much higher efficiencies than the current levels of less than 1%achieved by many crops. In this case the potential for improvement is enormous because most plants convert less than 1% of the solar energy falling on them to chemical energy through photosynthesis. Through genetic engineering, it is likely that this efficiency could be improved several-fold leading to vastly increased generation of biomass. Clearly, the achievement of sustainability employing high-level scientific developments will be an exciting development in decades to come

This page titled 2.5.1.6: Sustainable Energy- Away from the Sun and Back Again is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **1.6: Sustainable Energy- Away from the Sun and Back Again** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.1.7: Green Science

As shown in Figure 2.5.1.7.1, science, engineering, and technology are closely interrelated. Science refers to a set of rigorously organized bodies of knowledge and their acquisition based upon several criteria. Dealing with the discovery, explanation, and development of theories pertaining to interrelated natural phenomena of energy, matter, time, and space, science consists of an organized body of facts consistent with a number of general laws that are verifiable by rigorously defined systematic experimental processes composing the **scientific method**. Although science seeks to avoid value judgements, scientific methods are used for worthy goals, such as development of genetically based cures for disease, as well as for more sinister purposes, such as the synthesis of more deadly military poisons. **Pure science** is conducted to extend knowledge without defined practical goals and **applied science** is directed toward practical, usually commercial, objectives; these two aspects of science are commonly applied together. Especially with the explosive growth of the internet, which allows dissemination of both information and misinformation, it is important to beware of "junk science" often used to attempt to support political, economic, or theological agendas in areas such as climate change and evolution.

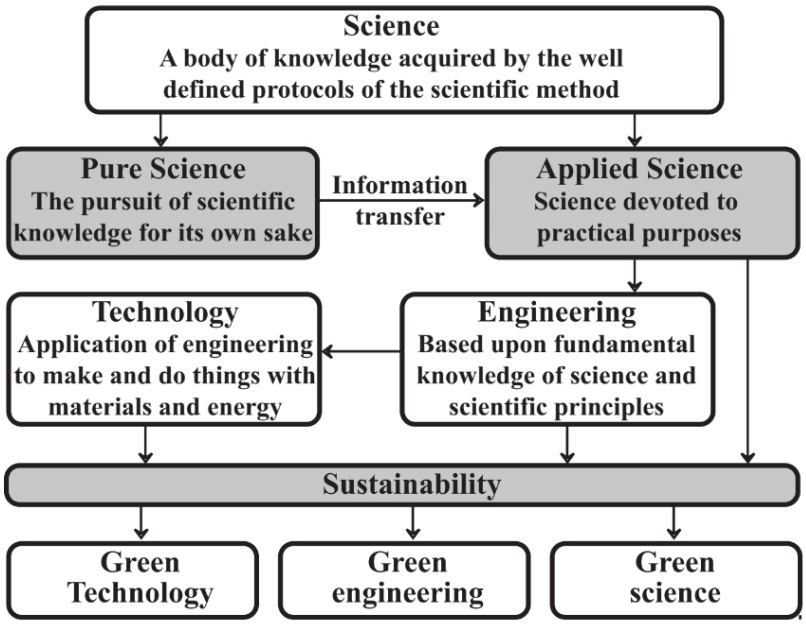


Figure 2.5.1.7.1, . Scientific principles are the basis of engineering which guides the development of technologies to make and dothings with materials and energy. As Earth's natural capital becomes more limiting, green technology, green engineering, and green science have become crucial to maintaining the planet's environment in a livable state.

The practice of science dedicated to sustainability including the maintenance of environmental quality, the reduction of hazards, the efficient use of environmentally benign sources of energy, and the minimization of consumption of non-renewable resources is the basis of **green science**, which is one of the main themes of this book. Chemical science, the science of matter, has great accomplishments to its credit in improving human welfare but it has also led to environmental pollution, exposure to hazardous substances, consumption of resources such as petroleum feedstocks, and other unpleasant aspects of modern industrialized societies. These problems can be mitigated by the constructive application of chemical science known as **green chemistry** defined as the *practice of chemistry in a manner that maximizes its benefits while eliminating or at least greatly reducing its adverse impacts.*[8] Based upon "twelve principles of green chemistry,"9since the mid-1990s the science of green chemistry has spawned a number of books, journal articles, and symposia as well as journals devoted to the topic and centers and societies of green chemistry. This book is primarily about green chemistry and how it relates to sustainability, green science, and green technology.

This page titled 2.5.1.7: Green Science is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **1.7: Green Science** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.1.8: Green Technology

Humans direct **technology** toward practical ends to make things they need with materials and to utilize energy in manufacturing, transportation, and the maintenance of hospitable living conditions. Long a matter of applied human ingenuity, technology is now mostly the product of engineering based on the fundamental knowledge of science and application of scientific principles. Technology uses the plans and means to achieve specific practical objectives provided by engineering to carry out desired objectives. Technology obviously has enormous importance in determining how human activities affect Earth and its life support systems.

Three great "growth spurts" in human populations that have taken place since modern humans first appeared on Earth have been enabled by developments in technology that enabled successively higher levels of human populations to exist. The first of these culminating about 10,000 years ago with a human population of perhaps 2 or 3 million was enabled by the primitive, but remarkably effective tools that early humans developed. One such technology was the bow and arrow which allowed for killing game at a much safer distance from the unwilling and often dangerous source of meat than that required for spearing or clubbing the quarry. The technology for making garments from animal hides allowed humans to avoid potentially fatal hypothermia from exposure to Ice Age climates.

About 10,000 years ago a second population growth spurt got underway as the "hunter/gatherer" societies that had sustained humans evolved into more reliable agricultural societies when humans learned to cultivate crops and domesticate animals used for meat, milk, and wool. These societies ensured a generally dependable food supply in smaller areas. As a result, humans were able to gather food from relatively small agricultural fields rather than having to scout large expanses of forest or grasslands for game to kill or berries to gather. Agricultural economic systems were based upon newly evolved technologies for cultivating soil, utilizing irrigation, and transporting food for trade by primitive sailing vessels. Ancillary technologies such as spinning wheels and looms for turning wool and plant fibers into cloth and water-powered mills for grinding grain also appeared. The development of agricultural systems had the major effect of allowing humans to remain in one place in settlements and, freed from the necessity of having to constantly hunt and gather food from their natural surroundings, humans could apply their ingenuity in areas such as developing more sophisticated tools. The agricultural revolution allowed a second large increase in numbers of humans and enabled the development of a human population of around 100 million 1000 years ago.

The third great growth spurt in human populations came with the industrial revolution, beginning slowly several centuries ago and made possible by the ability to utilize energy other than that provided by human labor and animal power. Initially wind power and water power were harnessed for mills and factories to use to produce goods. After about 1800 this power potential was multiplied many-fold with the steam engine and later the internal combustion engine, turbines, nuclear energy, and electricity, enabling current world population of almost 7 billion and growing at an eventually unsustainable rate (though not as fast as some of the more pessimistic projections from past years).

Unintended Consequences and the Need for Green Technology

According to the **law of unintended consequences**, the predicted benefits of new technologies are often accompanied by substantial problems, sometimes called revenge effects, due to the unforeseen ways in which people interact with new technologies. The individual freedom of movement and huge economic boost that resulted from the nascent automobile industry were accurately predicted by visionaries in the early 1900s, but they did not foresee the millions of deaths from automobile accidents, unhealthy polluted air in urban areas, urban sprawl, and depletion of petroleum resources that occurred in the following century. The tremendous educational potential of personal computers was visualized when the first of these came on the market. Less predictable were the mind-numbing hours that young people (and some not so young) would waste playing senseless computer games or viewing questionable content on the internet.

Defined as technology applied in a manner that minimizes environmental impact and resource consumption and maximizes economic output relative to materials and energy, **green technology** is designed to foresee and avoid revenge effects. Aided by increasingly sophisticated computer methodologies, the practitioners of green technology attempt to predict undesirable consequences of new technologies and put in place preventative measures before revenge effects have a chance to develop and cause major problems. A key component of the implementation of green technology is careful consideration of the practice of **industrial ecology**, which integrates the principles of science, engineering, and ecology in industrial systems through which goods

and services are provided in a way that minimizes environmental impact and optimizes utilization of resources, energy, and capital. Above all a sustainable means of providing goods and services, industrial ecology considers every aspect of so doing from concept, through production, and to the final fate of products remaining after use. It is most successful in its application when it mimics natural ecosystems, which are inherently sustainable by nature. Rather than organisms and populations of organisms working in natural ecosystems, industrial ecology works through **industrialecosystems** consisting of groups of industrial concerns, distributors, and other enterprises functioning to mutual advantage, using each others' products, recycling each others' potential waste materials, and utilizing energy as efficiently as possible. Industrial ecology is discussed in some detail in Chapter 13, "The Anthrosphere, Green Chemistry, and Industrial Ecology."

This page titled 2.5.1.8: Green Technology is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **1.8: Green Technology** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.1.9: Sustainability and the Eco-Economy

An important component of green technology is **life-cycle analysis** (assessment) which considers process and product design in the management of materials from their source through manufacturing, distribution, use, reuse (recycle), and ultimate fate. The objective of life-cycle analysis is to determine, quantify, and minimize adverse resource, environmental, economic, and social impacts. The four major facets of life-cycle analysis are(1)determination of the scope of the assessment; (2) inventory analysis of materials mass and energy to enable development of mass and energy balances; (3) analysis of impact on the environment, human health, and other impacted areas; and (4) improvement analysis to determine ways in which greater efficiencies maybe achieved.

Life-cycle analysis is summarized in Figure 2.5.1.9.1. Note that there are several possible recycling loops ranging from simple product reuse through material reprocessing and fabrication to waste mining in which wastes are processed to reclaim useful materials that can go back into the manufacturing process.

The **eco-economy** is one in which the production of goods and services is totally integrated with the natural world. The practice of **eco-efficiency** (Figure 2.5.1.9.2) enables provision of affordable goods and services to satisfy human needs sustainably, doing so with the minimum consumption of Earth's natural capital and with most efficient utilization of energy. The practice of eco-efficiency has several major aspects related to sustainability. Dematerialization seeks to meet economic needs with minimum amounts of material using renewable and recycled sources wherever possible. Analogous to dematerialization is "de-energyization" which uses only minimum amounts of energy and takes energy from renewable sources. Service and knowledge flows are substituted for material flows wherever possible. Using natural ecosystems as models, production loops are closed to the maximum extent possible. Sustainability is enhanced by shifting from a supply-driven to a demand-driven economy. Rather than manufacturing large quantities of material goods that are then vigorously marketed, the emphasis is placed upon finding the real needsof consumers, then meeting them in the most efficient way possible. Functional extension is achieved by manufacturing products with enhanced functionality and selling services to increase product functionality. Eco-efficient products are designed to be as durable as possible consistent with their intended uses, to be long-lived, and designed for ease of recycling of components and materials. Dispersion of toxic materials is minimized or eliminated in eco-efficient systems.

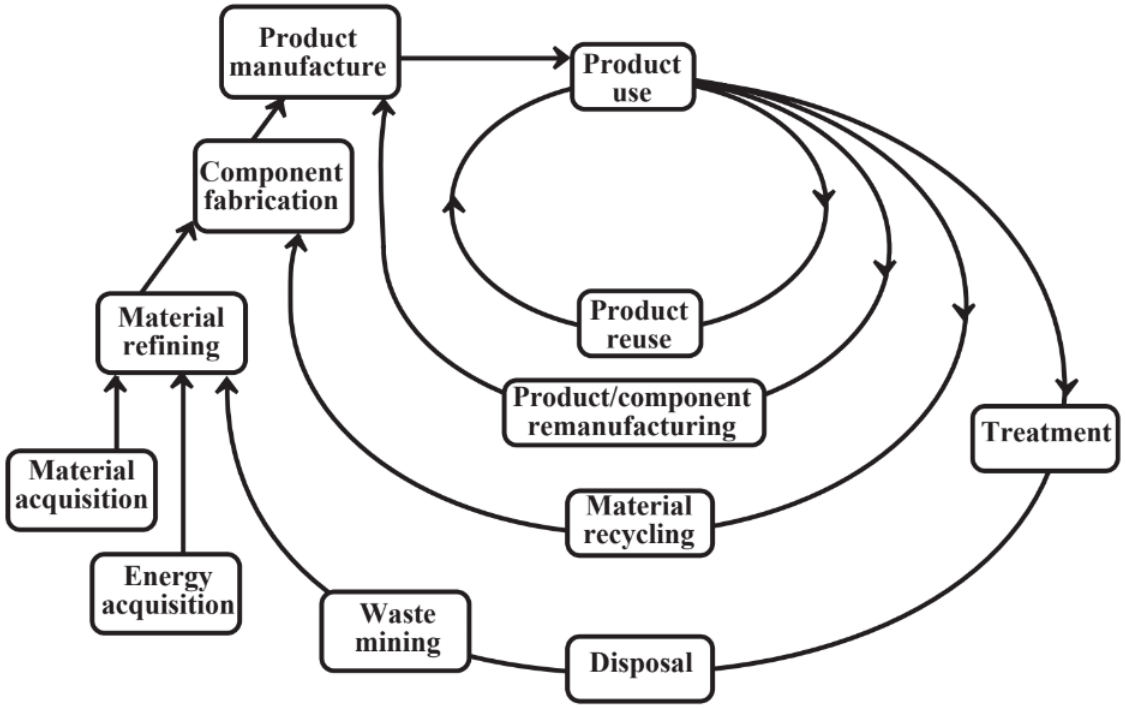
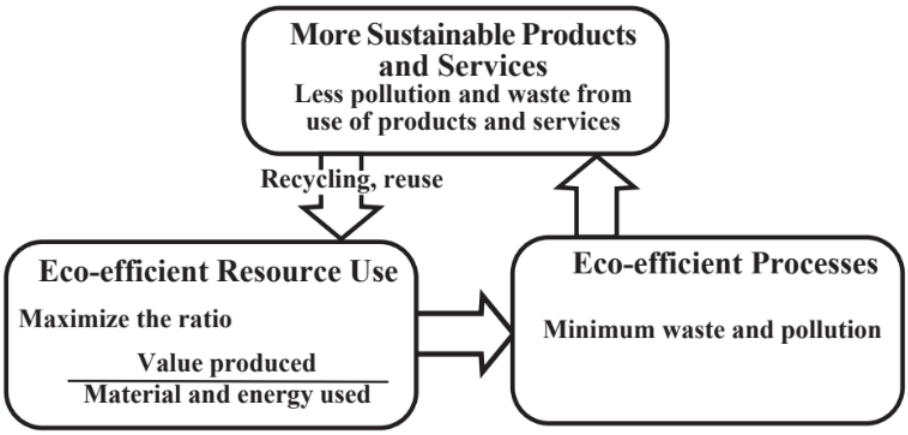


Figure 2.5.1.9.1. Various levels of materials use employed in life-cycle analysis. The inner loops are most desirable from the viewpoint of



sustainability

Figure 2.5.1.9.2
Achievement of sustainability with eco-efficient resource use and eco-efficient processes

---

This page titled 2.5.1.9: Sustainability and the Eco-Economy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **1.9: Sustainability and the Eco-Economy** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.1.10: Green Products and Services- Design for Sustainability

Green products use relatively smaller amounts of material and energy and throughout their life cycles from manufacture to disposal minimize exposure of humans and the environment to hazardous substances, pollutants, and wastes. A **greenservice** fulfills these criteria in providing a service. A hybrid fuel/electric automobile with a capability of recharging its battery from the electrical grid has minimum environmental and sustainability impact so it is a green product. The function of such an automobile can be replaced by a green service consisting of efficient rail and bus transportation.

Green products have several characteristics in common. One is high durability so long as it does not pose undue disposal problems. Another characteristic is low potential for exposure to toxic substances. A green product comes with minimal, recyclable packaging and is generally reusable, repairable, and capable of being remanufactured. Green materials used in consumer applications are relatively more concentrated with minimum inert ingredients so that they are economical to transport (a concentrated liquid laundry detergent compared to detergent in a granular formulation containing a lot of filler). Business and governmental practices and infrastructures can determine product sustainability. For example, a sustainable product requires that it be easily repaired and that repair parts are readily available. The sustainability of electrical batteries requires dropoff points and infrastructure for recycling of the materials in the batteries. A good example is the need for facilities to reclaim and recycle the lithium in lithium ion batteries as these sources of electrical energy storage become more popular in a world where sources of lithium are limited. Efforts have been made in some countries to require product take-back to enforce recycling components and materials in spent products.

The practices of design **for environment and design for sustainability are** key aspects of eco-efficiency.[10]In the practice of design for environment, environmental performance and potential environmental impacts are given priority consideration in the earliest stages of product development including raw materials acquisition, manufacturing, packaging, distribution, installation, operation, and ultimate fate at the end of the useful product lifetime. Design for sustainability does much the same thing with emphasis upon minimizing impact upon Earth's natural capital.

The following is a list of some specific design characteristics that go into design for environment and design for sustainability:

- Material substitution to use readily available materials from renewable sources that require relatively less energy for processing and that are recyclable,nontoxic, and environmentally friendly
- Minimize packaging
- Products with long lifetimes
- Promote recycling and reuse with products designed for separability, disassembly, reuse, remanufacture and recyclability.
- Consumptive materials should be biodegradable or capable of being burned for energy without emitting harmful byproducts (avoid bound halogens and heavy metals in plastics)

LITERATURE CITED

1. “World Commission on Environment and Development,”*Our Common Future*, Oxford University Press, New York, 1987.
2. Owen, David, Green Metropolis: Why Living Smaller, Living Closer, and Driving Less are the Keys to Sustainability, Riverhead Hardcover, 2009, New York.
3. Raven, Peter H., Linda R. Berg, and David M. Hassenzahl, *Environment*, 6th ed.,Wiley, 2008.
4. Carson, Rachel, *Silent Spring*, Fawcett Crest, New York, 1964.
5. “The Tragedy of the Commons,” Garrett Hardin, *Science*, **162**, 1243, (1968).
6. Duncan, Dayton,*Miles from Nowhere*, Viking, New York, 1994, p. 145.
7. Baum, Rudy M., “Sustainability: Learning to Live Off the Sun in Real Time,”*Chemical and Engineering News*, **86**, 42-46 (2008).
8. Anastas, Paul, Irvin J. Levy, and Kathryn E. Parent, *Green Chemistry Education: Changing the Course of Chemistry*, Oxford University Press, Oxford, UK, 2009.
9. Anastas, Paul T., and John C. Warner, *Green Chemistry: Theory and Practice*, Oxford University Press, Oxford, UK, 2009.
10. Jasch, Christine, *Environmental and Material Flow Cost Accounting: Principles and Procedures (Eco-Efficiency in Industry and Science)*, Springer, Berlin, 2008.

This page titled 2.5.1.10: Green Products and Services- Design for Sustainability is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **1.10: Green Products and Services- Design for Sustainability** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.1.11: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. According to some environmentalists, Earth is now entering the Anthropocene epoch. What is an epoch? In which epoch have humans lived until now? What are some of the past epochs that Earth has been through? Approximately when was the term Anthropocene coined and who is responsible for it?

2. It is noted in this chapter that the narrow layer between the geosphere and atmosphere where plants grow is an important environmental interface. The microclimate in this interface may vary significantly in its characteristics from the microclimate of the atmosphere just above it. What is microclimate? How does the microclimate in the narrow layer just described tend to vary from the climate just above it?

3. Thin layers and interfaces are very important in Earth's environment. Describe some of these layers.

4. The assertion has been made that if Earth were a classroom globe, the layer of soil covering it would be about as thick as the dimensions of a human cell. Assume that a classroom globe is 25 centimeters in diameter. Look up Earth's dimensions and using the assertion made above, calculate the average thickness of soil in cm.

5. What was the London smog disaster? When did it occur? What interaction between the atmosphere and anthrosphere caused it? Approximately how many people died as a result?

6. Approximately when did the steam engine become a practical machine? How did it enable the industrial revolution to occur? What are the analogies with current conditions in which increasingly large and sophisticated farm machinery are enabling an ongoing revolution in agriculture?

7. What was the Love Canal affair and how did it likely influence environmental laws and implementation of environmental regulations in the U.S.?

8. One of the most active areas in energy development in many nations including the U.S. is the development of new sources of natural gas. What is happening in the area of natural gas utilization? Does burning natural gas emit carbon dioxide to the atmosphere? How is it superior in that respect to coal and petroleum? What is meant by the term "bridging fuel" as applied to natural gas?

9. How are desertification and deforestation related to global warming? How do they contribute to each other? How does destruction of a rain forest contribute carbon dioxide to the atmosphere?

10. The process for chemical fixation of atmospheric nitrogen developed by Bosch and Haber in the early 1900s led to the ability to make huge quantities of explosives that subsequently took millions of lives in warfare. However, the initial product, NH_3, is not explosive. Look up the formulas of several common explosives and suggest what is done to use NH_3 to make explosives.

11. Secretary of Energy Steven Chu has suggested three areas in which Nobel-level breakthroughs are needed in the achievement of energy sustainability. Considering conditions on Earth and the rate of depletion of natural capital suggest other areas in the general area of sustainability in which breakthroughs are needed.

12. Of all nations, Brazil has been the most successful in using fuels from biological sources as renewable energy sources. What are the conditions in Brazil that have made that possible?

13. What are the major crops that enabled humans to transition from hunter/gatherer societies to agricultural societies? Where is it believed that this transition first took place and how long ago was it?

14. How does photochemical smog that plagues Los Angeles, Mexico City, and many other urban areas around the world illustrate the law of unintended consequences and revenge effects?

15. A basic premise of Green Science and Technology is that "human welfare must be measured interms of quality of life, not just acquisition of material possessions." Suggest how the dwellings of humans and their living surroundings in general might reflect

such a transition.

16. Figure 1.5 reflects various levels of materials use in which the innermost loops are most desirable. The most efficient materials use is product reuse in which a product or component is put directly back into the manufacturing loop. Suggest how a high level of component reuse at the manufacturing site might in fact reflect a less than optimum manufacturing process.

17. Gross domestic product per unit of energy use is a measure of efficiency of the economic systems of various countries and probably reflects the degree to which needed goods and services are provided relative to burden on natural capital. Look up the rank of various nations with respect to this ratio. Are there any surprises in this list? Are there nations on this list with a high rank that are probably not very desirable in terms of living amenities and are there others in which the opposite is true? Suggest two or three nations that have both a high rank and a high quality of life in general?

18. Much concern is being expressed over deteriorating infrastructure in the U.S. What is infrastructure and how may its degeneration contribute to a lack of sustainability?

19. Are paper grocery bags necessarily “green?” After doing some research on paper manufacture suggest ways in which they may not be ideal for sustainability. What is a greener alternative?

20. One recipe for sustainability may be expressed as “electrons, not paper.” Suggest what is meant by this expression and how it may be implemented for sustainability

---

This page titled 2.5.1.11: Questions and Problems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **Questions and Problems** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

2.5.2: The Key Role of Chemistry and Making Chemistry Green

“We have no choice but to deal with chemistry because all things are chemical. The human body itself is a remarkably complex chemical system composed of thousands of chemicals, the main one of which is water.”

2.5.2.1: Chemistry is Good (and Unavoidable)

2.5.2.2: The Environment and its Chemistry

2.5.2.3: What is Environmental Chemistry?

2.5.2.4: Environmental Pollution

2.5.2.5: Practice of Green Chemistry

2.5.2.6: Green Chemistry and Synthetic Chemistry

2.5.2.7: Reduction of Risk- Hazard and Exposure

2.5.2.8: The Risks of No Risks

2.5.2.9: Waste Prevention

2.5.2.10: Basic Principles of Green Chemistry

2.5.2.11: Some Things to Know About Chemistry before You Even Start

2.5.2.12: Combining Atoms to Make Molecules and Compounds

2.5.2.13: The process of making and breaking chemical bonds- chemical reactions

2.5.2.14: The Nature of Matter and States of Matter

2.5.2.15: References and Questions

This page titled 2.5.2: The Key Role of Chemistry and Making Chemistry Green is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.2.1: Chemistry is Good (and Unavoidable)

Chapter 1 has provided an overview of environmental science and sustainability in general. Sustainability is all about the way that we deal — often poorly and wastefully — with matter and energy. As the science of matter and how it interacts with energy, **chemistry** is a crucial science in sustainability. Most of this book is about the sustainable practice of chemistry.

Many people are freaked out by the idea of chemistry and try to avoid it. But avoiding chemistry is impossible. That is because all matter, all things, the air we must breathe, the water we must drink, and all living organisms are made of chemicals. People who try to avoid all things that they regard as chemical may fail to realize that chemical processes are continuously being carried out in their own bodies. These are processes that far surpass in complexity and variety those that occur in chemical manufacturing operations. So, even those people who want to do so cannot avoid chemistry; they are, themselves, complex chemical factories. The best course of action with anything that cannot be avoided and that might have an important influence on our lives (one's chemistry professor may come to mind) is to try to understand it, to deal with it. To gain an understanding of chemistry is probably one of the main reasons why you are reading this book.

As one of its major functions, this book seeks to present a body of chemical knowledge from the most fundamental level within a framework of the relationship of chemical science to human beings, their surroundings, and their environment. Face it, the study of chemistry-based upon facts about elements, atoms, compounds, molecules, chemical reactions, and other fundamental concepts needed to understand this science, though enticing to some, is found by many to be less than exciting. However, these concepts and many more are essential to a meaningful understanding of chemistry. Anyone interested in green chemistry clearly wants to know how chemistry influences people in the world around us. So this book discusses real-world chemistry, introducing chemical principles as needed.

During the approximately two centuries that chemical science has been practiced on an ever-increasing scale, it has enabled the production of a wide variety of goods that are valued by humans. These include such things as pharmaceuticals that have improved health and extended life, fertilizers that have greatly increased food productivity and prevented widespread starvation, and semiconductors that have made possible computers and other electronic devices. Without the persistent efforts of chemists and the enormous productivity of the chemical industry, nothing approaching the high standard of living enjoyed in modern industrialized societies would be possible.

But there can be no denying that in years past, and even at present, chemistry has been misused in many respects, such as the release of pollutants and toxic substances and the production of nonbiodegradable materials, resulting in harm to the environment and living things, including humans. It is now obvious that chemical science must be turned away from emphasis upon the exploitation of limited resources and the production of increasing amounts of products that ultimately end up as waste and toward the application of chemical science in ways that provide for human needs without damaging the Earth support system and depleting its natural capital (defined in Chapter 1) upon which all living things depend. Fortunately, the practice of chemical science and industry is moving steadily in the direction of environmental friendliness and resource sustainability. The practice of chemistry in a manner that maximizes its benefits while eliminating or at least greatly reducing its adverse impacts has come to be known as green chemistry, the central theme of this book.

As will be seen in later chapters of this book, the practice of chemistry is divided into several main categories. Most elements other than carbon are involved with **inorganic chemistry**. Common examples of inorganic chemicals are water, salt (sodium chloride), air pollutant sulfur dioxide, and lime. Carbon occupies a special place in chemistry because it is so versatile in the kinds of chemical species (compounds) that it forms. Most of the tens of millions of known chemicals are substances based on carbon. These compounds are organic chemicals and addressed by the subject of organic **chemistry**. The unique chemistry of carbon is addressed specifically in Chapter 6, "The Wonderful World of Carbon: Organic Chemistry and Biochemicals." The underlying theory and physical phenomena that explain chemical processes are explained by physical **chemistry**. Living organisms carry out a vast variety of chemical processes that are important in green chemistry and environmental chemistry. The chemistry that living organisms perform is biochemistry, which is addressed specifically in Chapters 7, "Chemistry of Life andGreen Chemistry." It is always important to know the identities and quantities of various chemical species present in a system, including various environmental systems. Often, significant quantities of chemical species are very low, so sophisticated means must be available to detect and quantify such species. The branch of chemistry dealing with the determination of kinds and quantities of chemical species is analytical **chemistry**.

As the chemical industry developed and grew during the early and mid-1900s, most practitioners of chemistry remained unconcerned with and largely ignorant of the potential for harm — particularly damage to the outside environment — of their products and processes. Environmental chemistry was essentially unknown and certainly not practiced by most chemists. Incidents of pollution and environmental damage, which were many and severe, were commonly accepted as a cost of doing business or blamed upon the industrial or commercial sectors. The unfortunate attitude that prevailed is summarized in a quote from a standard book on industrial chemistry from 1954 (*American Chemical Industry—A History*, W. Haynes, Van Nostrand Publishers, 1954): “By sensible definition any by-product of a chemical operation for which there is no profitable use is a waste. The most convenient, least expensive way of disposing of said waste — up the chimney or down the river — is best."

Despite their potential to cause harm, nobody is more qualified to accept responsibility for environmental damage from chemical products or processes than are chemists who have the knowledge to understand how such harmful effects came about. As the detrimental effects of chemical manufacture and use became more obvious and severe, chemists were forced, often reluctantly, to deal with them. At present, enlightened chemists and chemical engineers do not view the practice of environmentally beneficial chemistry and manufacturing as a burden, but rather as an opportunity that challenges human imagination and ingenuity.

This page titled 2.5.2.1: Chemistry is Good (and Unavoidable) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.1: Chemistry is Good (and Unavoidable)** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.2: The Environment and its Chemistry

It was noted in Section 1.2 that, the *environment* consists of our surroundings, which may affect us and which, in turn, we may affect. Obviously the chemical nature and processes of matter in the environment are important. Compared to the generally well defined processes that chemists study in the laboratory, those that occur in the environment are rather complex and must be viewed in terms of simplified models. A large part of this complexity is due to the fact that the environment consists of the five overlapping and interacting spheres — the atmosphere, the hydrosphere, the geosphere, the biosphere, and the anthrosphere mentioned in Section 1.2 and shown from the viewpoint of their chemical interactions in Figure 2.5.2.2.1. In order to better understand the chemistry that occurs in these spheres they are briefly described here.

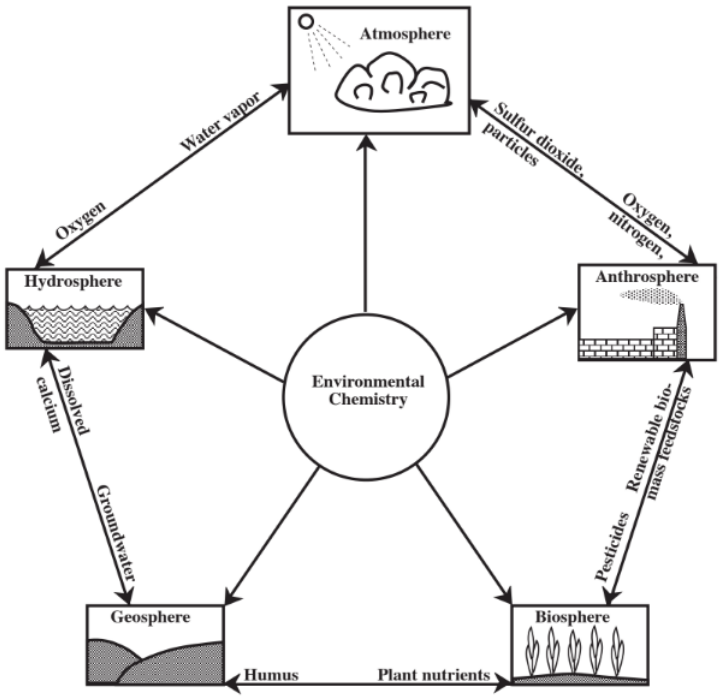


Figure 2.5.2.2.1. Illustration of the five major spheres of the environment and some of their chemical interactions. These spheres are closely tied together, interact with each other, and exchange materials and energy. A meaningful examination of environmental sciences must include all five of these spheres, including the anthrosphere.

The atmosphere is a very thin layer compared to the size of Earth, with most atmospheric gases lying within a few kilometers of sea level. In addition to providing oxygen for living organisms, the atmosphere provides carbon dioxide required for plant photosynthesis, and nitrogen that organisms use to make proteins. The atmosphere serves a vital protective function in that it absorbs highly energetic ultraviolet radiation from the sun that would kill living organisms exposed to it. A particularly important part of the atmosphere in this respect is the stratospheric layer of ozone, an ultraviolet-absorbing form of elemental oxygen. Because of its ability to absorb infrared radiation by which Earth loses the energy that it absorbs from the sun, the atmosphere stabilizes Earth's surface temperature. The atmosphere also serves as the medium by which the solar energy that falls with greatest intensity in equatorial regions is redistributed away from the Equator. It is the medium in which water vapor evaporated from oceans as the first step in the hydrologic cycle is transported over land masses to fall as rain over land.

Earth's water is contained in the hydrosphere. Although frequent reports of torrential rainstorms and flooded rivers produced by massive storms might give the impression that a large fraction of Earth's water is fresh water, more than 97% of it is seawater in the oceans. Most of the remaining fresh water is present as ice in polar ice caps and glaciers. A small fraction of the total water is present as vapor in the atmosphere. The remaining liquid fresh water is that available for growing plants and other organisms and for industrial uses. This water may be present on the surface as lakes, reservoirs, and streams, or it may be underground as groundwater.

The solid part of earth, the geosphere, includes all rocks and minerals. A particularly important part of the geosphere is soil, which supports plant growth, the basis of food for all living organisms. The **lithosphere** is a relatively thin solid layer extending from

Earth's surface to depths of 50–100 km. The even thinner outer skin of the lithosphere known as the **crust** is composed of relatively lighter silicate-based minerals. It is the part of the geosphere that is available to interact with the other environmental spheres and that is accessible to humans.

The biosphere is composed of all living organisms. For the most part, these organisms live on the surface of the geosphere on soil, or just below the soil surface. The oceans and other bodies of water support high populations of organisms. Some life forms exist at considerable depths on ocean floors. In general, though, the biosphere is a very thin layer at the interface of the geosphere with the atmosphere. The biosphere is involved with the geosphere, hydrosphere, and atmosphere in *biogeochemicalcycles* through which materials such as nitrogen and carbon are circulated.

Through human activities, the anthrosphere, that part of the environment made and operated by humans, has developed strong interactions with the other environmental spheres. Many examples of these interactions could be cited. By cultivating large areas of soil for domestic crops, humans modify the geosphere and influence the kinds of organisms in the biosphere. Humans divert water from its natural flow, use it, sometimes contaminate it, then return it to the hydrosphere. Emissions of particles to the atmosphere by human activities affect visibility and other characteristics of the atmosphere. The emission of large quantities of carbon dioxide to the atmosphere by combustion of fossil fuels may be modifying the heat-absorbing characteristics of the atmosphere to the extent that global warming is almost certainly taking place. The anthrosphere perturbs various biogeochemical cycles.

The effect of the anthrosphere over the last two centuries in areas such as burning large quantities of fossil fuels is especially pronounced upon the atmosphere and has the potential to change the nature of Earth significantly. According to Nobel Laureate Paul J. Crutzen of the Max Planck Institute for Chemistry, Mainz, Germany, this impact is so great that it will lead to a new global epoch to replace the halocene epoch that has been in effect for the last 10,000 years since the last Ice Age. Dr. Crutzen has coined the term *anthropocene* (from anthropogenic) to describe the new epoch that is upon us.

This page titled 2.5.2.2: The Environment and its Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.2: The Environment and its Chemistry** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.3: What is Environmental Chemistry?

The practice of green chemistry must be based upon **environmental chemistry**. This important branch of chemical science is defined as *the study of the sources, reactions, transport, effects, and fates of chemical species in water, soil, air, and living environments and the effects of technology thereon.*[1] Figure 2.2 illustrates this definition of environmental chemistry with an important type of environmental chemical species. In this example, two of the ingredients required for the formation of photochemical smog — nitric oxide and hydrocarbons — are emitted to the atmosphere from vehicles and transported through the atmosphere by wind and air currents. In the atmosphere, energy from sunlight brings about photochemical reactions that convert nitric oxide and hydrocarbons to ozone, noxious organic compounds, and particulate matter, all characteristic of photochemical smog. Various harmful effects are manifested, such as visibility-obscuring particles in the atmosphere, or ozone, which is unhealthy when inhaled by humans, or toxic to plants. Finally, the smog products end up on soil, deposited on plant surfaces, or in bodies of water.

Figure 2.2.1 showing the five environmental spheres may provide an idea of the complexity of environmental chemistry as a discipline. Enormous quantities of materials and energy are continually exchanged among the five environmental spheres. In addition to variable flows of materials, there are variations in temperature, intensity of solar radiation, mixing, and other factors, all of which strongly influence chemical conditions and behavior.

Throughout this book the role of environmental chemistry in the practice of green chemistry is emphasized. Green chemistry is practiced to minimize the impact of chemicals and chemical processes upon humans, other living organisms, and the environment as a whole. It is only within the framework of a knowledge of environmental chemistry that green chemistry can be successfully practiced.

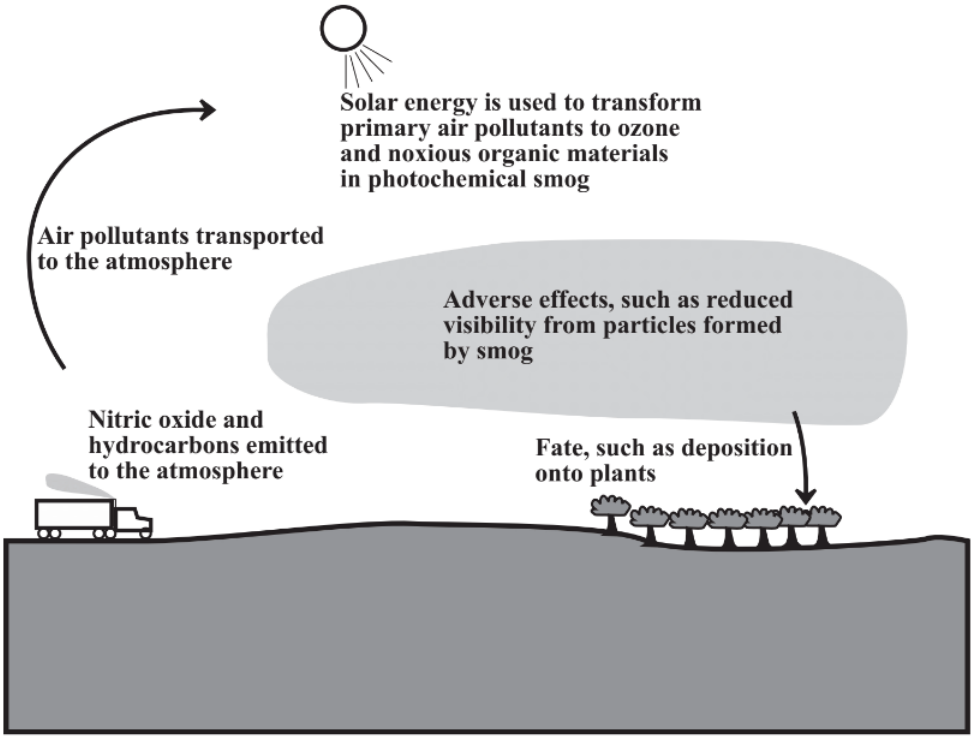


Figure 2.5.2.3.1 Illustration of the definition of environmental chemistry with a common environmental contaminant

There are several highly interconnected and overlapping categories of environmental chemistry. **Aquatic chemistry** deals with chemical phenomena and processes in water. Aquatic chemical processes are very strongly influenced by microorganisms in the water, so there is a strong connection between the hydrosphere and biosphere insofar as such processes are concerned. Aquatic chemical processes occur largely in “natural waters” consisting of water in oceans, bodies of fresh water, streams, and underground aquifers. These are places in which the hydrosphere can interact with the geosphere, biosphere, and atmosphere and is often

subjected to anthrospheric influences. Aspects of aquatic chemistry are considered in various parts of this book and are addressed specifically in Chapter 9, “Water, the Ultimate Green Substance.”

Atmospheric chemistry is the branch of environmental chemistry that considers chemical phenomena in the atmosphere. Two things that make this chemistry unique are the extreme dilution of important atmospheric chemicals and the influence of photochemistry. Photochemistry occurs when molecules absorb photons of high-energy visible light or ultraviolet radiation, become energized (“excited”), and undergo reactions that lead to a variety of products, such as photochemical smog. In addition to reactions that occur in the gas phase, many important atmospheric chemical phenomena take place on the surfaces of very small solid particles suspended in the atmosphere and in droplets of liquid in the atmosphere. Although no significant atmospheric chemical reactions are mediated by organisms in the atmosphere, microorganisms play a strong role in determining species that get into the atmosphere. As examples, bacteria growing in the absence of oxygen, such as in cows’ stomachs and under water in rice paddies, are the single greatest source of hydrocarbon in the atmosphere because of the large amounts of methane that they emit. The greatest source of organic sulfur compounds in the atmosphere consists of microorganisms in the oceans that emit dimethyl sulfide. Atmospheric chemistry is addressed specifically in Chapter 10, “Blue Skies for a Green Environment.”

Chemical processes that occur in the geosphere involving minerals and their interactions with water, air, and living organisms are addressed by the topic of **geochemistry**. A special branch of geochemistry, soil chemistry, deals with the chemical and biochemical processes that occur in soil. Aspects of geochemistry are explained in Chapter 11, “The Geosphere and a Green Earth,” and soil and agricultural chemistry are covered in Chapter 12, “The Biosphere and Feeding a Hungry World.

Environmental biochemistry addresses biologically mediated processes that occur in the environment. Such processes include, as examples, the biodegradation of organic waste materials in soil or water and processes within biogeochemical cycles, such as denitrification, which returns chemically bound nitrogen to the atmosphere as nitrogen gas. The basics of biochemistry are presented in Chapter 7, “The Chemistry of Life and Green Chemistry,” and other aspects of biochemistry are presented in Chapter 12, “The Biosphere and Feeding a Hungry World.” Chapter 14, “Feeding the Anthrosphere: Utilizing Renewable and Biological Materials,” discusses how chemical processes carried out by organisms can produce material feedstocks needed for the practice of green chemistry. The toxic effects of chemicals are of utmost concern to chemists and the public. Chapter 16, “Terrorism, Toxicity, And Vulnerability: Green Chemistry and Technology in Defense of Human Welfare,” deals with aspects of these toxic effects and discusses **toxicological chemistry.**

Although there is not a formally recognized area of chemistry known as “anthrospheric chemistry,” most of chemical science and engineering developed to date deals with chemistry carried out in the anthrosphere. Included is industrial chemistry, which is very closely tied to the practice of green chemistry. A good way to view “anthrospheric chemistry” from a green chemistry perspective is within the context of **industrial ecology**. Industrial ecology considers industrial systems in a manner analogous to natural ecosystems. In a system of industrial ecology, various manufacturing and processing operations carry out “industrial metabolism” on materials.A successful industrial ecosystem is well balanced and diverse, with various enterprises that generate products for each other and use each other’s products and potential wastes. A well-functioning industrial ecosystem recycles materials to the maximum extent possible and produces little — ideally no — wastes. Therefore, a good industrial ecosystem is a green chemical system. Industrial ecology and anthrospheric environmental chemistry are addressed in Chapter 13, “TheAnthrosphere, Green Chemistry, and Industrial Ecology."

This page titled 2.5.2.3: What is Environmental Chemistry? is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.3: What is Environmental Chemistry?** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.4: Environmental Pollution

Environmental chemistry has developed in response to problems and concerns regarding environmental pollution. Although awareness of chemical pollution had increased significantly in the two decades following World War II, the modern environmental movement dates from the 1962 publication of Rachel Carson's classic book *Silent Spring*. The main theme of this book was the accumulation of DDT and other persistent, mostly pesticidal, chemicals through the food chain, which caused birds at the end of the chain to produce eggs with soft shells that failed to produce viable baby birds. The implication was that substances harming bird populations might harm humans as well.

Around the time of the publication of Silent Spring another tragedy caused great concern regarding the potential effects of chemicals. This was the occurrence of approximately 10,000 births of children with badly deformed or missing limbs as a result of their mothers having taken the pharmaceutical thalidomide to alleviate the effects of morning sickness at an early stage of pregnancy.

The 1960s were a decade of high concern and significant legislative action in the environmental arena aimed particularly at the control of water and air pollutants. By around 1970, it had become evident that the improper disposal of chemicals to the geosphere was also a matter of significant concern. Although many incidents of such disposal were revealed, the one that really brought the problem into sharp focus was the Love Canal site in Niagara Falls, New York. This waste dump was constructed in an old abandoned canal in which large quantities of approximately 80 waste chemicals had been placed for about two decades starting in the 1930s. It had been sealed with a clay cap and given to the city. A school had been built on the site and housing constructed around it. By 1971 it became obvious that the discarded chemicals were leaking through the cap. This problem led eventually to the expenditure of many millions of dollars to remediate the site and to buy out and relocate approximately one thousand households. More than any other single incident the Love Canal problem was responsible for the passage of legislation in the U.S., including Superfund, to clean up hazardous waste sites and to prevent their production in the future.

By about 1970 it was generally recognized that pollution of air, water, and land was reaching intolerable levels. As a result, various countries passed and implemented laws designed to reduce pollutants and to clean up waste chemical sites at a cost that has easily exceeded one trillion dollars globally. In many respects, this investment has been strikingly successful. Streams that had deteriorated to little more than stinking waste drainage ditches (the Cuyahoga River in Cleveland, Ohio, once caught on fire from petroleum waste floating on its surface) have been restored to a healthy and productive condition. Despite a much increased population, the air quality in smog-prone Southern California has improved markedly. A number of dangerous waste disposal sites have been cleaned up. Human exposure to toxic substances in the workplace, in the environment, and in consumer products has been greatly reduced. The measures taken and regulations put in place have prevented devastating environmental problems from occurring.

Initially, serious efforts to control pollution were based on a **command and control** approach, which specifies maximum concentration guideline levels of substances that can be allowed in the atmosphere or water and places limits on the amounts or concentrations of pollutants that can be discharged in waste streams. Command and control efforts to diminish pollution have resulted in implementation of various technologies to remove or neutralize pollutants in potential waste streams and stack gases. These are so-called end-of-pipe measures. As a result, numerous techniques, such as chemical precipitation of water pollutants, neutralization of acidic pollutants, stack gas scrubbing, and waste immobilization have been developed and refined to deal with pollutants after they are produced.

Release of chemicals to the environment is now tracked in the U.S. through the Toxics Release Inventory TRI, under requirements of the Emergency Planning and Community Right to Know Act, which requires that information be provided regarding the release of more than 300 chemicals. The release of as much as one billion kilograms of these chemicals has been reported in the U.S. during a single year. Not surprisingly, the chemical industry produces the most such substances, followed by primary metals and paper manufacture. Significant amounts are emitted from transportation equipment, plastics, and fabricated metals, with smaller quantities from a variety of other enterprises. Although the quantities of chemicals released are high, they are decreasing, and the publicity resulting from the required publication of these releases has been a major factor in decreasing the amounts of chemicals released.

Although much maligned, various pollution control measures implemented in response to command and control regulation have reduced wastes and improved environmental quality. Regulation-based pollution control has clearly been a success and well worth the expense and effort. However, it is much better to prevent the production of pollutants rather than having to deal with them after they are made. This was recognized in United States with the passage of the 1990 Pollution Prevention Act, which recognized that source reduction is fundamentally different and more desirable than waste management and pollution control. This act explicitly states that, wherever possible, wastes are not to be generated and their quantities are to be minimized. The means for accomplishing this objective can range from very simple measures, such as careful inventory control and reduction of solvent losses due to evaporation, to much more sophisticated and drastic approaches including complete redesign of manufacturing processes with waste minimization as a top priority. The means for preventing pollution are best implemented through the practice of green chemistry, which is discussed in detail in the following section.

This page titled 2.5.2.4: Environmental Pollution is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.4: Environmental Pollution** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.5: Practice of Green Chemistry

The limitations of a command and control system for environmental protection have become more obvious even as the system has become more successful. In industrialized societies with good, well-enforced regulations, most of the easy and inexpensive measures that can be taken to reduce environmental pollution and exposure to harmful chemicals have been implemented. Therefore, small increases in environmental protection now require relatively large investments in money and effort. Is there a better way? There is, indeed. The better way is through the practice of green chemistry.

Green chemistry can be defined as the practice of chemical science and manufacturing in a manner that is sustainable, safe, and non-polluting and that consumes minimum amounts of materials and energy while producing little or no waste material. This definition of green chemistry is illustrated in Figure 2.5.2.5.1. The practice of green chemistry begins with recognition that the production, processing, use, and eventual disposal of chemical products may cause harm when performed incorrectly. In accomplishing its objectives, green chemistry and green chemical engineering may modify or totally redesign chemical products and processes with the objective of minimizing wastes and the use or generation of particularly dangerous materials. Those who practice green chemistry recognize that they are responsible for any effects on the world that their chemicals or chemical processes may have. Far from being economically regressive and a drag on profits, green chemistry is about increasing profits and promoting innovation while protecting human health and the environment.

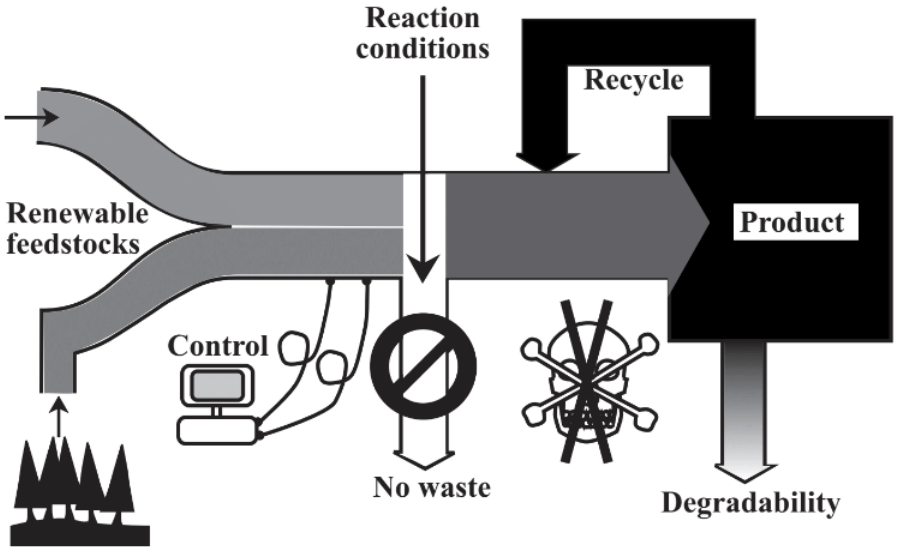


Figure 2.5.2.5.1. Illustration of the definition of green chemistry, which emphasizes renewable feedstocks, exacting control, mild reaction conditions, maximum recycling of materials, minimal wastes, and degradability of products that might enter the environment.

To a degree, we are still finding out what green chemistry is. That is because it is a rapidly evolving and developing subdiscipline in the field of chemistry. And it is a very exciting time for those who are practitioners of this developing science. Basically, green chemistry harnesses a vast body of chemical knowledge and applies it to the production, use, and ultimate disposal of chemicals in a way that minimizes consumption of materials, exposure of living organisms, including humans, to toxic substances, and damage to the environment. And it does so in a manner that is economically feasible and cost effective. In one sense, green chemistry is the most efficient possible practice of chemistry and the least costly when all of the costs of doing chemistry, including hazards and potential environmental damage are taken into account.

Green chemistry is sustainable chemistry. There are several important respects in which green chemistry is sustainable:

- Economic: At a high level of sophistication green chemistry normally costs less in strictly economic terms (to say nothing of environmental costs) than chemistry as it is normally practiced.
- Materials: By efficiently using materials, maximum recycling, and minimum use of virgin raw materials, green chemistry is sustainable with respect to materials.

- Waste: By reducing insofar as possible, or even totally eliminating their production, green chemistry is sustainable with respect to wastes

---

This page titled 2.5.2.5: Practice of Green Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.5: Practice of Green Chemistry** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.6: Green Chemistry and Synthetic Chemistry

Synthetic chemistry is the branch of chemical science involved with developing means of making new chemicals and developing improved ways of synthesizing existing chemicals. A key aspect of green chemistry is the involvement of synthetic chemists in the practice of environmental chemistry. Synthetic chemists, whose major objective has always been to make new substances and to make them cheaper and better, have come relatively late to the practice of environmental chemistry. Other areas of chemistry have been involved much longer in pollution prevention and environmental protection. From the beginning, analytical chemistry has been a key to discovering and monitoring the severity of pollution problems. Physical chemistry has played a strong role in explaining and modeling environmental chemical phenomena. The application of physical chemistry to atmospheric photochemical reactions has been especially useful in explaining and preventing harmful atmospheric chemical effects including photochemical smog formation and stratospheric ozone depletion. Other branches of chemistry have been instrumental in studying various environmental chemical phenomena. Now the time has arrived for the synthetic chemists, those who make chemicals and whose activities drive chemical processes, to become intimately involved in making the manufacture, use, and ultimate disposal of chemicals as environmentally friendly as possible.

Before environmental and health and safety issues gained their current prominence, the economic aspects of chemical manufacture and distribution were relatively simple and straightforward. The economic factors involved included costs of feedstock, energy requirements, and marketability of product. Now, however, costs must include those arising from regulatory compliance, liability, end-of-pipe waste treatment, and costs of waste disposal. By eliminating or greatly reducing the use of toxic or hazardous feedstocks and catalysts and the generation of dangerous intermediates and byproducts, green chemistry eliminates or greatly reduces the additional costs that have come to be associated with meeting environmental and safety requirements of conventional chemical manufacture.

As illustrated in Figure 2.5.2.6.1, there are two general and often complementary approaches to the implementation of green chemistry in chemical synthesis, both of which challenge the imaginations and ingenuity of chemists and chemical engineers. The first of these is to use existing feedstocks but make them by more environmentally benign, “greener,” processes. The second approach is to substitute other feedstocks that are made by environmentally friendly means. In some cases, a combination of the two approaches is used.

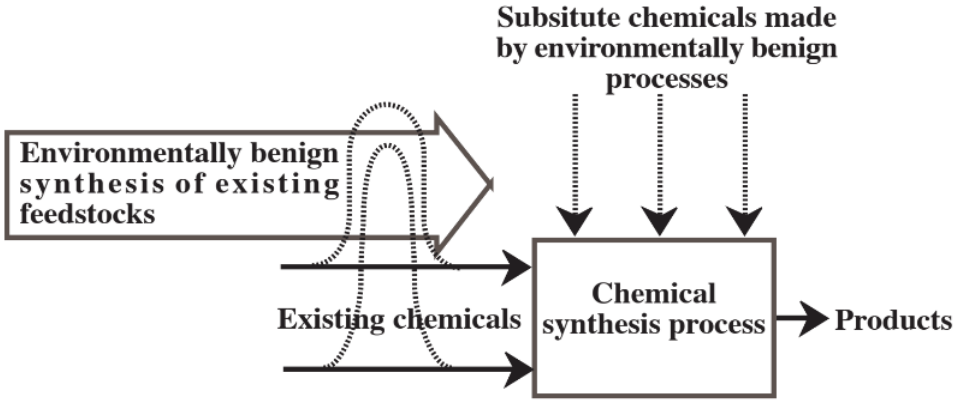


Figure 2.4. Two general approaches to the implementation of green chemistry. The dashed loops on the left represent alternative approaches to environmentally benign means of providing chemicals already used for chemical synthesis. A second approach, where applicable, is to substitute entirely different, environmentally safer raw materials.

Yield and Atom Economy

Traditionally, synthetic chemists have used **yield**, defined as a percentage of the degree to which a chemical reaction or synthesis goes to completion to measure the success of a chemical synthesis. For example, if a chemical reaction shows that 100 grams of product should be produced, but only 85 grams is produced, the yield is 85%. A synthesis with a high yield may still generate significant quantities of useless byproducts if the reaction does so as part of the synthesis process. Instead of yield, green chemistry emphasizes **atom economy**, the fraction of reactant material that actually ends up in final product. With 100 percent atom

economy, all of the material that goes into the synthesis process is incorporated into the product. For efficient utilization of raw materials, a 100% atom economy process is most desirable. Figure 2.7.1 illustrates the concepts of yield and atom economy.

This page titled 2.5.2.6: Green Chemistry and Synthetic Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

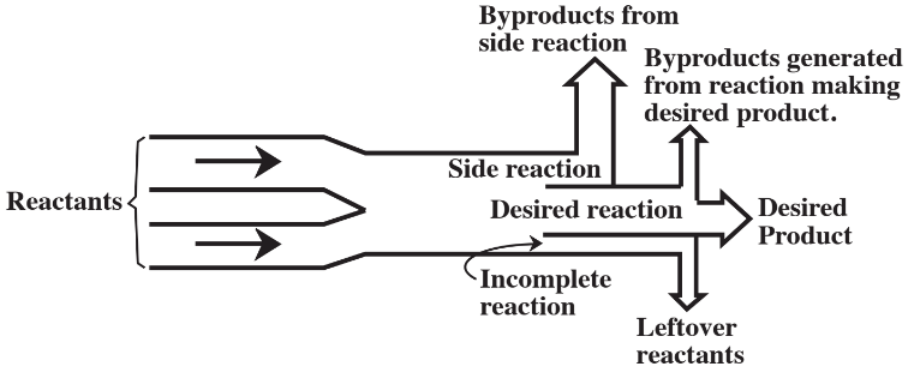
- **2.6: Green Chemistry and Synthetic Chemistry** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.7: Reduction of Risk- Hazard and Exposure

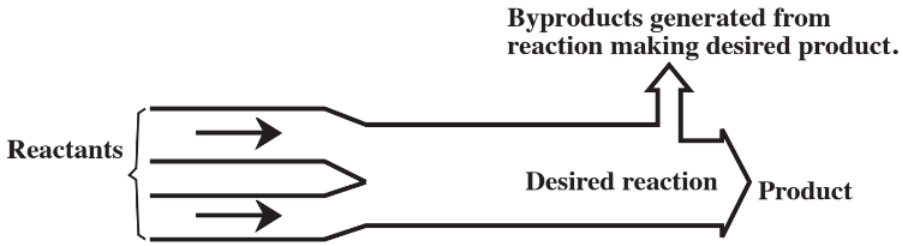
A major goal in the manufacture and use of commercial products, and, indeed, in practically all areas of human endeavor, is the reduction of risk. There are two major aspects of risk — the hazard presented by a product or process and exposure of humans or other potential targets to those hazards.

$$\text{Risk} = \text{Fhazard x exposure}$$

This relationship simply states that risk is a function of hazard times exposure. It shows that risk can be reduced by a reduction of hazard, a reduction of exposure, and various combinations of both.



(a) Typical reaction with less than 100% yield and with byproducts



(b) Reaction with 100% yield, but with byproducts inherent to the reaction



(c) Reaction with 100% atom economy, no leftover reactants, no byproducts

Figure 2.5.2.7.1 Illustration of percent yield and atom economy.

The command and control approach to reducing risk has concentrated upon reduction of exposure. Such efforts have used various kinds of controls and protective measures to limit exposure. The most common example of such a measure in the academic chemistry laboratory is the wearing of goggles to protect the eyes. Goggles will not by themselves prevent acid from splashing into the face of a student, but they do prevent the acid from contacting fragile eye tissue. Explosion shields will not prevent explosions, but they do retain glass fragments that might harm the chemist or others in the vicinity.

Reduction of exposure is unquestionably effective in preventing injury and harm. However, it does require constant vigilance and even nagging of personnel, as any laboratory instructor charged with making laboratory students wear their safety goggles at all times will attest. It does not protect the unprotected, such as a visitor who may walk bare-faced into a chemical laboratory ignoring the warnings for required eye protection. On a larger scale, protective measures may be very effective for workers in a chemical

manufacturing operation but useless to those outside the area or the environment beyond the plant walls who do not have protection. Protective measures are most effective against acute effects, but less so against long-term chronic exposures that may cause toxic responses over many years period of time. Finally, protective equipment can fail and there is always the possibility that humans will not use it properly.

Where feasible, hazard reduction is a much more certain way of reducing risk than is exposure reduction. The human factors that play so prominently in successfully limiting exposure and that require a conscious, constant effort are much less crucial when hazards have been reduced. Compare, for example, the use of a volatile, flammable, somewhat toxic organic solvent used for cleaning and degreasing of machined metal parts with that of a water solution of a nontoxic cleaning agent used for the same purpose. To safely work around the solvent requires an unceasing effort and constant vigilance to avoid such hazards as formation of explosive mixtures with air, presence of ignition sources that could result in a fire, and excessive exposure by inhalation or absorption through skin that might cause peripheral neuropathy (a nerve disorder) in workers. Failure of protective measures can result in a bad accident or serious harm to worker health. The water-based cleaning solution, however, would not present any of these hazards so that failure of protective measures would not create a problem.

Normally, measures taken to reduce risk by reducing exposure have an economic cost that cannot be reclaimed in lower production costs or enhanced value of product. Of course, failure to reduce exposure can have direct, high economic costs in areas such as higher claims for worker compensation. In contrast, hazard reduction often has the potential to substantially reduce operating costs. Safer feedstocks are often less costly as raw materials. The elimination of costly control measures can lower costs overall. Again, to use the comparison of an organic solvent compared to a water-based cleaning solution, the organic solvent is almost certain to cost more than the aqueous solution containing relatively low concentrations of detergents and other additives. Whereas the organic solvent will at least require purification for recycle and perhaps even expensive disposal as a hazardous waste, the water solution may be purified by relatively simple processes, and perhaps even biological treatment, then safely discharged as wastewater to a municipal wastewater treatment facility. It should be kept in mind, however, that not all low-hazard materials are cheap, and may be significantly more expensive than their more hazardous alternatives. And, in some cases, nonhazardous alternatives simply do not exist.

This page titled 2.5.2.7: Reduction of Risk- Hazard and Exposure is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.7: Reduction of Risk- Hazard and Exposure** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.8: The Risks of No Risks

There are limits to the reduction in risk beyond which efforts to do so become counterproductive. As in other areas of endeavor, there are circumstances in which there is no choice but to work with hazardous substances. Some things that are inherently dangerous are rendered safe by rigorous training, constant attention to potential hazards, and understanding of hazards and the best way to deal with them. Consider the analogy of commercial flight. When a large passenger aircraft lands, typically 100 tons of aluminum, steel, flammable fuel, and fragile human flesh traveling at a speed of twice the legal interstate speed limits for automobiles come into sudden contact through air-filled rubber tires with an unforgiving concrete runway. That procedure is inherently dangerous! But it is carried out millions of times per year throughout the world with but few injuries and fatalities, a tribute to the generally superb design, construction, and maintenance of aircraft and the excellent skills and training of aircrew. The same principles that make commercial air flight generally safe also apply to the handling of hazardous chemicals by properly trained personnel under carefully controlled conditions.

So, although much of this book is about risk reduction as it relates to chemistry, we must always be mindful of the risks of not taking risks. If we become so timid in all of our enterprises that we refuse to take risks, scientific and economic progress will stagnate. The U.S. space program is an example of an area in which progress has been made only by a willingness to take risks. However, progress has probably been slowed because of risk aversion resulting from previous accidents, especially the 1987 Challenger space shuttle tragedy. If we get to the point that no chemical can be made if its synthesis involves the use of a potentially toxic or otherwise hazardous substance, the progress of chemical science and the development of such beneficial products as new life-saving drugs or innovative chemicals for treating water pollutants may be held back. It may be argued that thermonuclear fusion entails significant risks as an energy source and that research on controlled thermonuclear fusion must therefore be stopped. But when that potential risk is balanced against the virtually certain risk of continuing to use fossil fuels that produce greenhouse gases that cause global climate warming, it seems sensible to at least continue research on controlled thermonuclear fusion energy sources. Another example is the use of thermal processes for treating hazardous wastes, somewhat risky because of the potential for the release of toxic substances or air pollutants, but still the best way to convert many kinds of hazardous wastes to innocuous materials.

This page titled 2.5.2.8: The Risks of No Risks is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.8: The Risks of No Risks** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.9: Waste Prevention

Waste prevention is better than having to treat or clean up wastes. In the earlier years of chemical manufacture the direct costs associated with producing large quantities of wastes were very low because such wastes were simply discarded into waterways, onto the ground, or in the air as stack emissions. With the passage and enforcement of environmental laws after about 1970, costs for waste treatment increased steadily. For example, General Electric announced in April, 2010, that it had spent $561 in the first phase of dredging and removing PCBs from deposits in Hudson River sediments produced decades earlier when the extremely persistent PCBs were discarded to the river as wastes from the company's manufacture of electrical equipment. The cost of the second phase of cleanup of these wastes was projected to be much more than the $561million already spent. The cleanup of pollutants including asbestos, dioxins, pesticide manufacture residues, perchlorate and mercury are costing various concerns hundreds of millions of dollars. The eventual cleanup costs from the 2010 BP Deepwater Horizon oil well blowout in the Gulf of Mexico may eventually exceed the $20 billion that the company initially set aside for the project. By the year 2000 in the United States, costs of complying with environmental and occupational health regulations had grown to a magnitude similar to that of research and development for industry as a whole. From a purely economic standpoint, therefore, a green chemistry approach that avoids these costs is very attractive, in addition to its large environmental benefits.

Although the costs of such things as engineering controls, regulatory compliance, personnel protection, wastewater treatment, and safe disposal of hazardous solid wastes have certainly been worthwhile for society and the environment, they have become a large fraction of the overall expense of doing business. Companies must now do **full cost accounting**, taking into account the total costs of emissions, waste disposal, cleanup, and protection of personnel and the environment, none of the proceeds of which go into the final product.

This page titled 2.5.2.9: Waste Prevention is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.9: Waste Prevention** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.10: Basic Principles of Green Chemistry

From the preceding discussion, it should be obvious that there are certain basic principles of green chemistry. Some publications recognize "the twelve principles of green chemistry."[2] This section addresses the main ones of these.

As anyone who has ever spilled the contents of a food container onto the floor well knows, it is better to not make a mess than to clean it up once made. As applied to green chemistry, this basic rule means that *waste prevention is much better than waste cleanup*. Failure to follow this simple rule has resulted in most of the troublesome hazardous waste sites that are causing problems throughout the world today.

One of the most effective ways to prevent generation of wastes is to make sure that insofar as possible *all materials involved in making a product should be incorporated into the final product*. Therefore, the practice of green chemistry is largely about incorporation of all raw materials into the product, if at all possible. We would not likely favor a food recipe that generated a lot of inedible byproduct. The same idea applies to chemical processes. In that respect, the concept of atom economy discussed in Section 2.6 is a key component of green chemistry.

The use or generation of substances that pose hazards to humans and the environment should be avoided. Such substances include toxic chemicals that may be hazardous to workers. They include substances that are likely to become air or water pollutants and harm the environment or organisms in the environment. Here the connection between green chemistry and environmental chemistry is especially strong.

Chemical products should be as effective as possible for their designated purpose, but with minimum toxicity. The practice of green chemistry is making substantial progress in designing chemicals and new approaches to the use of chemicals such that effectiveness is retained and even enhanced while toxicity is reduced.

Chemical synthesis as well as many manufacturing operations make use of auxiliary substances that are not part of the final product. In chemical synthesis, such a substance consists of solvents in which chemical reactions are carried out. Another example consists of separating agents that enable separation of product from other materials. Since these kinds of materials may end up as wastes or (in the case of some toxic solvents) pose health hazards, *the use of auxiliary substances should be minimized and preferably totally avoided.*

Energy consumption poses economic and environmental costs in virtually all synthesis and manufacturing processes. In a broader sense, the extraction of energy, such as fossil fuels pumped from or dug out of the ground, has significant potential to damage the environment. Therefore, *energy requirements should be minimized*. One way in which this can be done is through the use of processes that occur near ambient conditions, rather than at elevated temperature or pressure. One successful approach to this has been the use of biological processes, which, because of the conditions under which organisms grow, must occur at moderate temperatures and in the absence of toxic substances. Such processes are discussed further in Chapters 13 and 14.

Raw materials extracted from earth are depleting in that there is a finite supply that cannot be replenished after they are used. So, wherever possible, *renewable raw materials should be used instead of depletable feedstocks*. As discussed further in Chapter 14, biomass feedstocks are highly favored in those applications for which they work. For depleting feedstocks, recycling should be practiced to the maximum extent possible.

In the synthesis of an organic compound (see Chapter 6), it is often necessary to modify or protect groups on the organic molecule during the course of the synthesis. This often results in the generation of byproducts not incorporated into the final product, such as occurs when a protecting group is bonded to a specific location on a molecule, then removed when protection of the group is no longer needed. Since these processes generate byproducts that may require disposal, *the use of protecting groups in synthesizing chemicals should be avoided insofar as possible.*

Reagents should be as selective as possible for their specific function. In chemical language, this is sometimes expressed as a preference for selective catalytic reagents over nonselective stoichiometric reagents.

Products that must be dispersed into the environment should be designed to break down rapidly into innocuous products. One of the oldest, but still one of the best, examples of this is the modification of the surfactant in household detergents 15 or 20 years after they were introduced for widespread consumption to yield a product that is biodegradable. The poorly biodegradable

surfactant initially used caused severe problems of foaming in wastewater treatment plants and contamination of water supplies. Synthesis of a biodegradable substitute solved the problem.

Exacting “real-time” control of chemical processes is essential for efficient, safe operation with minimum production of wastes. This goal has been made much more attainable by modern computerized controls. However, it requires accurate knowledge of the concentrations of materials in the system measured on a continuous basis. Therefore, *the successful practice of green chemistry requires real-time, in-process monitoring techniques coupled with process control.*

Accidents, such as spills, explosions, and fires, are major hazards in the chemical industry. Not only are these incidents potentially dangerous in their own right, they tend to spread toxic substances into the environment and increase exposure of humans and other organisms to these substances. For this reason, it is best to *avoid the use or generation of substances that are likely to react violently, burn, build up excessive pressures, or otherwise cause unforeseen incidents in the manufacturing process*.

The principles outlined above are developed to a greater degree in the remainder of the book. They should be kept in mind in covering later sections.

This page titled 2.5.2.10: Basic Principles of Green Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.10: Basic Principles of Green Chemistry** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.11: Some Things to Know About Chemistry before You Even Start

Chapters 3-7 explain the basic principles of chemistry as they relate to green chemistry. For even greater detail on the basics of chemistry the reader is referred to a book on that subject.3However, at this point, it is useful to have a brief overview of chemistry, in a sense a minicourse in chemistry that provides the basic definitions and concepts of chemistry such as chemical compounds, chemical formulas, and chemical reactions before they are covered in detail in the later chapters.

All chemicals are composed of fewer than 100 naturally-occurring fundamental kinds of matter called elements. Humans have succeeded in making about 30 artificial elements since the late 1930s, but the amounts of these are insignificant compared to the total of known chemicals. Elements, in turn, are composed of very small entities called atoms. Atoms of the same element may differ a bit in their masses, but all atoms of the same element behave the same chemically. So we can logically begin the study of chemistry with the atoms that make up the elements of which all matter is composed.

Each atom of a particular element is chemically identical to every other atom. Each element is given an atomic number specific to the element, ranging from 1 to more than 100. The atomic number of an element is equal to the number of extremely small, positively charged protons contained in the nucleus located in the center of each atom of the element. Each electrically neutral atom has the same number of electrons as it has protons. The electrons are negatively charged and are in rapid motion around the nucleus, constituting a cloud of negative charge that makes up most of the volume of the atom. In addition to its atomic number, each element has a name and a chemical symbol, such as carbon, C; potassium, K (for its Latin name kalium); or cadmium, Cd. In addition to atomic number, name, and chemical symbol, each element has an atomic mass (atomic weight). The atomic mass of each element is the average mass of all atoms of the element so it is not a whole number. Atoms of most elements consist of two or more **isotopes**. All isotopes of the same element have identical chemical properties but differ in mass because of the presence in their nuclei of differing numbers of electrically neutral neutrons (see Chapter 3).

This page titled 2.5.2.11: Some Things to Know About Chemistry before You Even Start is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.11: Some Things to Know About Chemistry before You Even Start** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.12: Combining Atoms to Make Molecules and Compounds

About the only atoms that exist alone are those of the noble gases, a group of elements including helium, neon, argon, and radon located on the far right of the periodic table. Even the simple hydrogen atom in the elemental state is joined together with another hydrogen atom. Two or more uncharged atoms bonded together are called a **molecule**. As illustrated in Figure 2.5.2.12.1, the hydrogen molecule consists of 2 hydrogen atoms as denoted by the chemical formula of elemental hydrogen, H_2. This formula states that a molecule of elemental hydrogen consists of 2 atoms of hydrogen, shown by the subscript of 2. The atoms are joined together by a chemical bond. As explained in Chapter 3, a hydrogen atom consists of a very small positively charged nucleus surrounded by a much larger cloud of negative charge from a single, rapidly moving, electron. But, hydrogen atoms are more "content" with 2 electrons. So two hydrogen atoms share their two electrons constituting the chemical bond in the hydrogen molecule. A bond composed of shared electrons is a **covalent bond**.

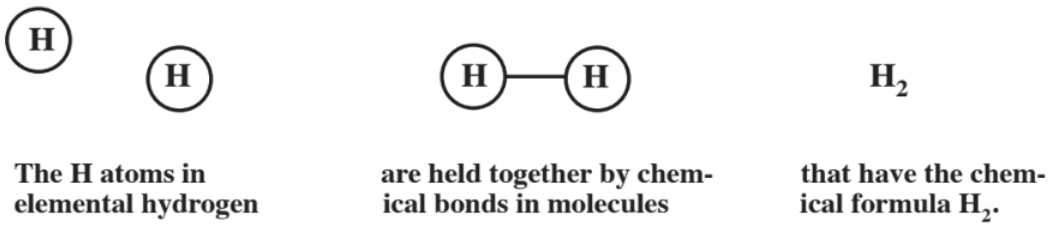


Figure 2.5.2.12.1. Molecule of H_2

Chemical Compounds

The example just discussed was one in which atoms of the same element, hydrogen, join together to form a molecule. Most molecules consist of atoms of different elements joined together. An example of such a molecule is that of **water**, chemical formula H_2O. This formula stands for the fact that the water molecule consists of *two* hydrogen atoms bonded to *one* oxygen atom, O, where the absence of a subscript number after the O indicates that there is 1 oxygen atom. The water molecule is shown in Figure 2.5.2.12.2. Each of the hydrogen atoms is held to the oxygen atom in the water molecule by two shared electrons in a covalent bond. A material such as water in which two or more elements are bonded together is called a **chemical compound**. It is because of the enormous number of combinations of two or more atoms of different elements that it is possible to make 20 million or more chemical compounds from fewer than 100 elements.

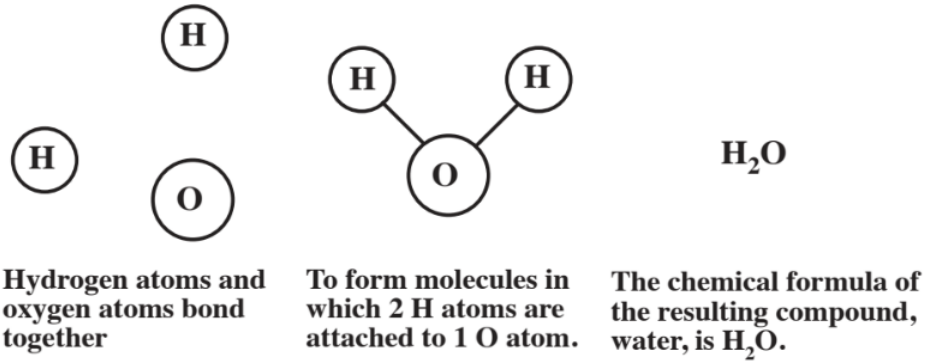


Figure 2.5.2.12.2. Atoms of hydrogen and oxygen join together by chemical bonds to form a molecule of water, chemical formula H_2O

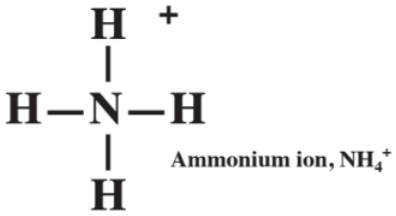
Ionic Bonds

Two different molecules have just been discussed in which atoms are joined together by *covalentbonds* consisting of *shared* electrons. Another way in which atoms can be joined together is by *transfer* of electrons from one atom to another. A single *neutral* atom has a number of electrons surrounding its nucleus that is the same as the number of protons in the nucleus of the atom. But, if the atom loses one or more negatively charged electrons, it ends up with a net *positive* electrical charge and the atom becomes a positively charged **cation**. An atom that has gained one or more negatively charged electrons attains a net *negative* charge and is

called an **anion**. Cations and anions are attracted together in an **ionic compound** because of their opposite electrical charges. The oppositely charged ions are joined by **ionicbonds** in a **crystalline lattice**.

Figure 2.5.2.12.3 shows the best known ionic compound, sodium chloride, NaCl (common table salt). The chemical formula of NaCl implies that there is 1 Na for each Cl. In this case these consist of Na^+ cations and Cl^- anions. For ionic compounds such as NaCl, the first part of the name is simply that of the metal forming the cation, in this case sodium. The second part of the name is based upon the anion, but has the ending *ide*. So the ionic compound formed from sodium and chlorine is sodium chlor*ide*. As shown by the preceding example, ionic compounds may consist of ions composed of atoms that have lost electrons (producing positively charged cations) and other atoms that have gained electrons (producing negatively charged anions).

In addition to being charged atoms, ions may also consist of groups of several atoms with a net charge. Ammonium ion, NH_4^+, is such an ion. As shown below, the NH_4^+ cation consists of 4H atoms covalently bonded (by 2 shared electrons) to a central N atom, with the group of 5 total atoms having a net electrical charge of +1.





The transfer of a single negatively charged electron from a neutral sodium atom (Na) to a neutral chlorine atom (Cl)

produces a positively charged sodium cation and a negatively charged chloride anion.

Sodium cations held together with equal numbers of chloride anions by the attraction of their opposite charges (ionic bonds) constitute the ionic compound NaCl.

Figure 2.5.2.12.3. Ionic bonds are formed by the transfer of electrons and the mutual attraction of oppositely charged ions in a crystalline lattice.

This page titled 2.5.2.12: Combining Atoms to Make Molecules and Compounds is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.12: Combining Atoms to Make Molecules and Compounds** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.13: The process of making and breaking chemical bonds- chemical reactions

The preceding section has discussed chemical compounds and the two major kinds of bonds— covalent bonds and ionic bonds — that hold them together. Next is discussed the process of making and taking apart chemical compounds, **chemical reactions**. A chemical reaction occurs when chemical bonds are broken and formed and atoms are exchanged to produce chemically different species.

First consider two very simple chemical reactions involving only one element, oxygen. In the very thin air high in the stratosphere more than 10 kilometers above Earth's surface (above the altitudes where jet airliners normally cruise), high-energy ultraviolet radiation from the sun, represented by the symbol $h\nu$, splits apart molecules of elemental oxygen, O_2,

$$O_2 + h\nu \rightarrow 2O$$

to produce oxygen atoms. As with most single atoms, the O atoms are reactive and combine with oxygen molecules to produce ozone, O_3:

$$O + O_2 \rightarrow O_3$$

Both of these processes are chemical reactions. In a chemical reaction, the substances on the left of the arrow (read as "yields") are the **reactants** and those on the right of the arrow are **products**. The first of these reactions states that the chemical bond holding together a molecule of O_2 *reactant* is split apart by the high energy of the ultraviolet radiation to produce two oxygen atom *products*. In the second reaction, an oxygen atom reactant, O, and an oxygen molecule reactant, O_2, form a chemical bond to yield an ozone product, O3. Are these very simple chemical reactions important to us? Emphatically yes. They produce a shield of ozone molecules in the stratosphere which in turn absorb ultraviolet radiation that otherwise would reach Earth's surface, destroying life, causing skin cancer and other maladies that would make our existence on Earth impossible. As discussed in Chapter 10, the use of chlorofluorocarbon refrigerants (Freons) has seriously threatened the stratospheric ozone layer. It is a triumph of environmental chemistry that this threat was realized in time to do something about it and an accomplishment of green chemistry to develop relatively safe substitutes for ozone-threatening chemicals.

Many chemical reactions are discussed in this book. At this point a very common chemical reaction can be considered, that of elemental hydrogen with elemental oxygen to produce water. A first approach to writing this reaction is

$$H_2 + O_2 \rightarrow H_2O$$

stating that elemental hydrogen and elemental oxygen react together to produce water. This is not yet a proper *chemicalequation* because it is not balanced. A **balanced chemical equation** has the same number of each kind of atom on both sides of the equation. As shown above, there are 2 H atoms in the single H_2 molecule on the left and 2 H atoms in the single molecule H_2O product. That balances hydrogen, but leaves 2 O atoms in the O_2 molecule on the left with only 1 O atom in the single H_2O molecule product. But, writing the reaction a

$$2H_2 + O_2 \rightarrow 2H_2O$$

gives a balanced chemical equation with a total of 4 H atoms in 2 H_2 molecules on the left, 4 H atoms in 2 H_2O molecules on the right, and a total of 2 O atoms in the 2 H_2O molecules on the right, which balances the 2 O atoms in the O_2 molecule on the left. So the equation as now written is *balanced*. A **balanced** chemical equation always has the same number of each kind of atom on both sides of the equation.

This page titled 2.5.2.13: The process of making and breaking chemical bonds- chemical reactions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.13: The process of making and breaking chemical bonds- chemical reactions** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.14: The Nature of Matter and States of Matter

We are familiar with matter in different forms. We live in an atmosphere of gas that is mostly N_2 with about 1/4 as much oxygen, O_2, by volume. We only become aware of the gas in the atmosphere when something is wrong with it, such as contamination by irritating air pollutants. A person stepping into an atmosphere of pure N_2 would not notice anything wrong immediately, but would die within a few minutes, not because N_2 is toxic, but because the atmosphere lacks life-giving oxygen. The same atmosphere that we breathe contains water in the gas form as water vapor. And we are also familiar, of course, with liquid water and with solid ice.

he air that we breathe, like most substances, is a **mixture** consisting of two or more substances. Air is a **homogeneousmixture** meaning that the molecules of air are mixed together at a molecular level. There is no way that we can take air apart by simple mechanical means, such as looking at it under a magnifying glass and picking out its individual constituents. Another common substance that is a homogeneous mixture is drinking water, which is mostly H_2O molecules, but which also contains dissolved O_2 and N_2 from air, dissolved calcium ions (Ca^{2+}), chlorine added for disinfection, and other materials.

A **heterogeneousmixture** is one that contains discernible and distinct particles that, in principle at least, can be taken apart mechanically. Concrete is a heterogeneous mixture. Careful examination of a piece of broken concrete shows that it contains particles of sand and rock embedded in solidified Portland cement.

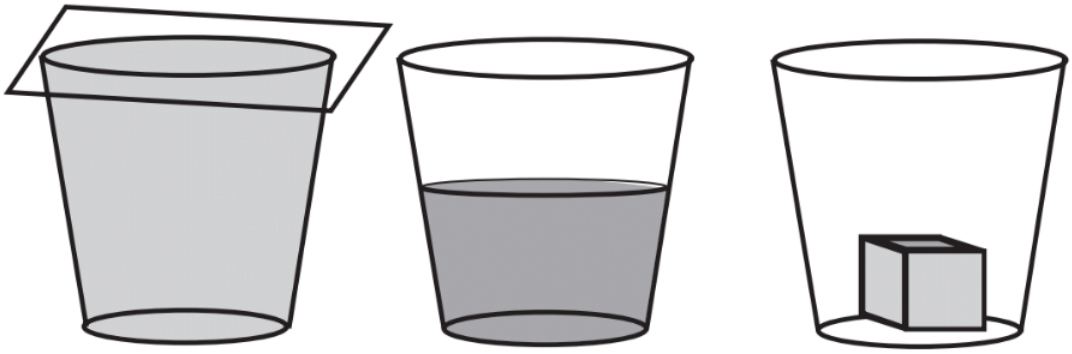
A material that consists of only one kind of substance is known as a **pure substance**. Absolutely pure substances are almost impossible to attain. Hyperpure water involved in semiconductor manufacturing operations approaches absolute purity. Another example is 99.9995% pure helium gas used in a combination gas chromatograph/mass spectrometer instrument employed for the chemical analysis of air and water pollutants.

Mixtures are very important in the practice of green chemistry. Among other reasons why this is so is that separation of impurities from mixtures in the processing of raw materials and in recycling materials is often one of the most troublesome and expensive aspects of materials utilization and may generate large quantities of wastes. Impurities may make mixtures toxic. For example, toxic arsenic, which is directly below phosphorus in the periodic table and has chemical properties similar to phosphorus, occurs as an impurity in the phosphate ores from which elemental phosphorus is extracted. This is not a problem for phosphorus used as fertilizer because the small amount of arsenic added to the soil is negligible compared to the arsenic naturally present in the soil. But, if the phosphorus is to be made into phosphoric acid and phosphate salts to be added to soft drinks or to food, impurity arsenic cannot be tolerated because of its toxicity requiring removal of this element at considerable expense.

Many byproducts of manufacturing operations are mixtures. For example, organochlorine solvents used to clean and degrease machined parts are mixtures that contain grease and other impurities. As part of the process for recycling these solvents, the impurities must be removed by expensive processes such as distillation. The separation of mixture constituents is often one of the most expensive aspects of the recycling of materials.

States of Matter

As shown in Figure 2.5.2.14.1, the three common states of matter are gases, liquids, and solids. These are readily illustrated by water, the most familiar form of which is liquid water. Ice is a solid and water vapor in the atmosphere or in a steam line is a gas.



Gas confined in a container has the same shape and volume of the container in which it is held

A quantity of liquid in a container assumes the shape of that part of the container in which it is held, but has a definite volume

The shape and volume of a solid are independent of the container in which it is placed

Figure 2.5.2.14.1

Illustration of the three states of matter.

Gases, such as those composing the air around us, are composed mostly of empty space through which molecules of the matter composing the gas move constantly, bouncing off each other or the container walls millions of times per second. A quantity of gas expands to fill the container in which it is placed. Because they are mostly empty space, gases can be significantly **compressed**; squeeze a gas and it responds with a decreased volume. Gas **temperature** is basically an expression of the tendency of the gas molecules to move more rapidly; higher temperatures mean faster molecular movement and more molecules bouncing off each other or container walls per second. The constant impact of gas molecules on container walls is the cause of gas **pressure**. Because of the free movement of molecules relative to each other and the presence of mostly empty space, a quantity of gas takes on the volume and shape of the container in which it is placed. The physical behavior of gases is described by several gas laws relating volumes of gas to quantities of the gas, pressure, and temperature. Calculations involving these laws are covered at the beginning of Chapter 10.

Molecules of **liquids** can move relative to each other, but cannot be squeezed together to a significant extent, so liquids are not compressible. Liquids do take on the shape of the part of a container that they occupy. Molecules of **solids** occupy fixed positions relative to each other. Therefore, solids cannot be significantly compressed and a solid object retains its shape regardless of the container in which it is placed.

This page titled 2.5.2.14: The Nature of Matter and States of Matter is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **2.14: The Nature of Matter and States of Matter** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.2.15: References and Questions

LITERATURE CITED

1. Manahan, Stanley E., *Environmental Chemistry*, 9th ed., Taylor & Francis/CRC Press, Boca Raton, FL, 2010.

2.Anastas, Paul T., and John C. Warner, *Green Chemistry Theory and Practice*, Oxford University Press, 1998.

3. Manahan, Stanley E., *Fundamentals of Sustainable Chemical Science*, Taylor & Francis/CRC Press, Boca Raton, FL, 2009.

SUPPLEMENTARY REFERENCES

Allen, David T., and David R. Shonnard, *Green Engineering: Environmentally Conscious Design of Chemical Processes*, Prentice Hall, Upper Saddle River, NH, 2002.

Brown, Lawrence S., and Thomas A. Holme, *Chemistry for Engineering Students*, 2nd ed., Brooks/Cole Cengage Learning, Belmont, CA, 2011.

Brown, Theodore L., H. Eugene Lemay, Bruce E. Bursten, Catherine J. Murphy, and Patrick Woodward, Chemistry: The Central Science, 11th ed., Prentice Hall, Upper Saddle River, NJ, 2008.

Anastas, Paul, Ed., *Handbook of Green Chemistry*, Wiley-VCH, New York, 2010.

Clark James and Duncan MacQuarrie, *Handbook of Green Chemistry and Technology*, Blackwell Science, Malden, MA, 2002.

Denniston, Katherine J., Joseph J. Topping, and Robert L. Caret, *General, Organic, and Biochemistry*, 7th ed., McGraw-Hill, New York, 2011.

Doble, Mukesh, and Anil Kumar Kruthiventi, *Green Chemistry and Processes*, Elsevier, Amsterdam, 2007.

Frost, Laura D., S. Todd Deal, and Karen C. Timberlake,*General, Organic, and Biological Chemistry: An Integrated Approach*, Prentice Hall, Upper Saddle River, NJ, 2010.

Guinn, Denise, and Rebecca Brewer, *Essentials of General, Organic, and Biochemistry: An Integrated Approach*, W. H. Freeman and Co., New York, 2009.

Hill, John W., and Doris Kolb, *Chemistry for Changing Times*, 11th ed, Prentice Hall, Upper Saddle River, NJ, 2006.

Horvath, Istvan T., and Paul T. Anastas, "Innovations and Green Chemistry," *Chemical Reviews*, **107**, 2169-2173 (2007).

Lancaster, Mike, *Green Chemistry*, Royal Society of Chemistry, London, 2002.

Li, Chao-Jun, and Barry M. Trost, "Green Chemistry for Chemical Synthesis," *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 13197-13202 (2008).

Matlack, Albert, *Introduction to Green Chemistry*, 2nd ed., CRC Press, Boca Raton, FL, 2010.

McMurry, John, David S. Ballantine, Carl A. Hoeger, Virginia E. Peterson, and Mary E. Castellion, *Fundamentals of General, Organic, and Biological Chemistry*, 6th ed., Prentice-Hall, Upper Saddle River, NJ, 2009.

Mickulecky, Peter J., Katherine Brutlag, Michelle Gilman, and Brian Peterson, *Chemistry Workbook for Dummies*, Wiley, Hoboken, NJ, 2008.

Middlecamp, Catherine, Steve Keller, Karen Anderson, Anne Bentley, Michael Cann, *Chemistry in Context*, 7th ed., McGraw-Hill, Dubuque, IA, 2011.

Roesky, Herbert W., Dietmar Kennepohl, and Jean-Marie Lehn, *Experiments in Green and Sustainable Chemistry*, Wiley-VCH, Weinheim, Germany, 2009.

Moore, John W., Conrad L. Stanitski, and Peter C. Jurs, *Chemistry: The Molecular Science*, 4th ed., Brooks/Cole Cengage Learning, Belmont, CA, 2011.

Smith, Janice G.,*Principles of General, Organic, and Biochemistry*, McGraw-Hill, New York, 2011.

Viegas, Jennifer, Ed., *Critical Perspectives on Planet Earth*, Rosen Publishing Group, New York, 2007.

QUESTIONS AND PROBLEMS

Access to and use of the internet are assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. What is chemistry? Why is it impossible to avoid chemistry?

2. What is green chemistry?

3. Match the following pertaining to major areas of chemistry:

    A. Analytical chemistry 1. Occurs in living organisms

    B. Organic chemistry 2. Underlying theory and physical phenomena

    C. Biochemistry 3. Chemistry of most elements other than carbon

    D. Physical chemistry 4. Chemistry of most carbon-containing compounds

    E. Inorganic Chemistry 5. Measurement of kinds and quantities of chemicals

4. What are the five environmental spheres? Which of these did not exist before humans evolved on Earth?

5. Discuss why you think the very thin "skin" of Earth ranging from perhaps two or three kilometers in depth below the surface to several kilometers (several miles) in altitude in the atmosphere has particular environmental importance.

6. What is environmental chemistry?

7. Which event may be regarded as the beginning of the modern environmental movement?

8. What is the command and control approach to pollution control?

9. What is the Toxics Release Inventory, TRI. How does it reduce pollution?

10. Why are incremental increases in regulations under the command and control approach to pollution control much less effective now than they were when pollution control laws were first enacted and enforced?

11. What is the special relationship of green chemistry to synthetic chemistry?

12. What does Figure 2.1 show with respect to environmental chemistry, green chemistry, and other topics discussed in this chapter?

13. In which important respects is green chemistry sustainable chemistry?

14. With respect to raw materials, what are two general and often complementary approaches to the practice of green chemistry?

15. What is the distinction between yield and atom economy?

16. For a chemical synthesis of a pharmaceutical compound, 100 kilograms (kg) of reactants mixed in the exact proportions that would give a 100% theoretical yield of product would give 65.2 kg of product, the rest being byproduct. In an actual chemical synthesis, an excess of 10kg of one of the reactants was added to the 100 kg of mixture to help push the reaction to completion. After the total of 110 kg of reactants was put through the process, an actual yield of 59.5 kg of product was obtained. What was the percent yield? What was the percent atom economy?

17. What are two factors that go into assessing risk?

18. What are the risks of no risks?

19. What are the major basic principles of green chemistry?

20. What is shown by the formula O_3? What about H_2O_2?

21. How does a covalent bond differ from an ionic bond?

22. What is the name given to a kind of material in which two or more different elements are bonded together?

23. Considering the compound shown in Figure 2.8, what is the name of the compound formed when a magnesium atom transfers two electrons to an oxygen atom giving a compound consisting of Mg^{2+} cations and O^{2-} anions?

24. Summarize the information given by $3H_2 + O_3 \rightarrow 3H_2O$.

25. In addition to showing the correct reactants and products, a correct chemical equation must be_____.

26. Name three kinds of matter based upon purity. Which of these is extremely rare?

27. In terms of molecules, how are gases, liquids, and solids distinguished?

28. Describe gas pressure and temperature in terms of molecular motion.

29.What is the Presidential Green Chemistry Challenge? What have been some of the winning ideas in this challenge?

30. Which elemental species mentioned in this chapter is present in photochemical smog?

This page titled 2.5.2.15: References and Questions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **References and Questions** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

2.5.3: The Elements - Basic Building Blocks of Green Chemicals

“More than fifty million unique chemical substances have been identified and the number is growing at a rapid pace. Essentially all of these are composed of the 92 naturally-occurring elements and the vast majority are from twenty or fewer of the most abundant elements.”

2.5.3.1: Elements, Atoms, and Atomic Theory

2.5.3.2: Hydrogen - The Simplest Atom

2.5.3.3: Helium - The First Noble Gas

2.5.3.4: Lithium, The First Metal

2.5.3.5: The Second Period of the Periodic Table

2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons

2.5.3.7: Completing the 20-Element Periodic Table

2.5.3.8: The Brief Periodic Table is Complete

2.5.3.E: The Elements - Basic Building Blocks of Green Chemicals (Exercises)

2.5.3.R: References

Template:HideTOC

This page titled 2.5.3: The Elements - Basic Building Blocks of Green Chemicals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.3.1: Elements, Atoms, and Atomic Theory

Chemistry is the science of matter. The fundamental building blocks of matter are the atoms of the various elements, which are composed of **subatomic particles**, the positively charged **proton** (+), the negatively charged **electron** (-), and the electrically neutral **neutron** (n). It is the properties of these atoms that determine matter's chemical behavior. More specifically, it is the arrangement and energy levels of electrons in atoms that direct how they interact with each other, thus dictating all chemical behavior. One of the most fundamental aspects of chemistry is that elemental behavior varies periodically with increasing atomic number. This has enabled placement of elements in an orderly arrangement with increasing atomic number known as the **periodic table**. The periodic behavior of elements' chemical properties is due to the fact that, as atomic number increases, electrons are added incrementally to atoms and occupy so-called shells, each filled with a specific number of electrons. After each shell is filled, a new shell is started, thus beginning a new period (row) of the periodic table. This sounds complicated, and indeed may be so, occupying the full-time computational activities of banks of computers to explain the behavior of electrons in matter. However, this behavior can be viewed in simplified models and is most easily understood for the first 20 elements using dots to represent electrons, enabling construction of an abbreviated 20-element periodic table. Although simple, this table helps understand and explain most of the chemical phenomena discussed in this book.

The chapter also emphasizes some of the green aspects of the first 20 elements and how they relate to sustainability. Included among these elements are the nitrogen, oxygen, carbon (contained in carbon dioxide), and hydrogen and oxygen (in water vapor) that make up most of the air in the "green" atmosphere; the hydrogen and oxygen in water, arguably the greenest compound of all; the sodium and chlorine in common table salt; the silicon, calcium, and oxygen that compose most mineral matter, including the soil that grows plants supplying food to most organisms; and the hydrogen, oxygen, carbon, nitrogen, phosphorus, and sulfur that are the predominant elements in all living material.

Long Before Subatomic Particles Were Known, There Was Dalton's Atomic Theory

Atomic theory describes the atoms in relation to chemical behavior. With the sophisticated tools now available to chemists, the nature of atoms, largely based upon the subatomic particles of which they are composed, especially the negatively charged electrons, is well known. But long before these sophisticated tools were even dreamed about, more than two centuries ago in 1808, an English schoolteacher named John Dalton came up with the atomic theory that bears his name. To a large extent, this theory is the conceptual basis of modern chemistry. Key aspects of Dalton's atomic theory are the following:

- The matter in each element is composed of extremely small particles called atoms. (Dalton regarded atoms as indivisible, unchanging bodies. We now know that they exchange and share electrons, which is the basis of chemical bonding.)
- Atoms of different elements have different chemical properties. (These differences may range from very slight, such as those between the noble gases neon and argon, to vastly different, such as those between highly metallic sodium and strongly nonmetallic chlorine.)
- Atoms cannot be created, destroyed, or changed to atoms of other elements. (In modern times, the provision is added that these things do not happen in ordinary chemical processes, since atoms can be changed to atoms of other elements by nuclear reactions, such as those that occur in nuclear reactors.)
- Chemical compounds are formed by the combination of atoms of different elements in definite, constant ratios that usually can be expressed as integers or simple fractions.
- Chemical reactions involve the separation and combination of atoms. (This phenomenon was surmised before anything was known about the nature of chemical bonds that are broken and formed as part of the process of chemical reactions.)

Three Important Laws

Dalton's atomic theory explains the three important laws listed below. Evidence for these laws had been found prior to the publishing of Dalton's atomic theory, and the atomic theory is largely based upon them.

1. **Law of Conservation of Mass**: *There is no detectable change in mass in an ordinary chemical reaction*. (This law, which was first stated in 1798 by "the father of chemistry," the Frenchman Antoine Lavoisier, follows from the fact that in ordinary chemical reactions no atoms are lost, gained, or changed; in chemical reactions, mass is conserved.)
2. **Law of Constant** Composition: *A specific chemical compound always contains the same elements in the same proportions by mass*.

3. **Law of Multiple Proportions**: *When two elements combine to form two or more compounds, the masses of one combining with a fixed mass of the other are in ratios of small whole numbers*. A common illustration of this law is provided by the simple hydrocarbon compounds of carbon and hydrogen, which include CH_4, C_2H_2, C_2H_4, and C_2H_6. In these compounds the relative masses of C and H are in ratios of small whole numbers.

The Nature of Atoms

At this point it is useful to note several characteristics of atoms, which were introduced in Section 2.11. Atoms are extremely small and extremely light. Their individual masses are expressed by the minuscule atomic mass unit, u. The sizes of atoms are commonly expressed in picometers, where a picometer is 0.000 000 001 millimeters (a millimeter is the smallest division on the metric side of a ruler). Atoms may be regarded as spheres with diameters between 100 and 300 picometers.

As noted at the beginning of this chapter, atoms are composed of three basic *subatomic particles*, the positively charged *proton*, the electrically neutral *neutron*, and the much lighter negatively charged *electron*. Each proton and neutron has a mass of essentially 1 atomic mass unit, whereas the mass of the electron is only about 1/2000 as much The protons and neutrons are located in the *nucleus* at the center of the atom and the electrons compose a “fuzzy cloud” of negative charge around the nucleus. Essentially all the mass of an atom is in the nucleus and virtually all the volume is in the cloud of electrons. Each atom of a specific element has the same number of protons in its nucleus. This is the *atomic number* of the element. Each element has a name and is represented by a **chemical symbol** consisting of one or two letters. Atoms of the same element that have different numbers of neutrons and, therefore, different masses, are called **isotopes**. Isotopes may be represented by symbols such as ${}^{12}_{6}C$ where the subscript is the atomic number and the superscript is the **mass number**, which is the sum of the numbers of protons and neutrons in an atom.

The average mass of all the atoms of an element is the **atomic mass**. Atomic masses are expressed relative to the carbon ${}^{12}_{6}C$ the isotope, which contains 6 protons and 6 neutrons in its nucleus. The mass of this isotope is taken as exactly 12 u. Atomic masses normally are not integers, in part because atoms of most elements consist of two or more isotopes with different masses.

Electrons in Atoms

The behavior of electrons in the cloud of negative charge making up most of the volume of atoms, particularly their energy levels and orientations in space, are what determine chemical behavior. Arrangements of electrons are described by *electron configuration*. A detailed description of electron configuration is highly mathematical and sophisticated, but is represented in a very simplified fashion in this chapter. Because of their opposite charges, electrons are strongly attracted to positively charged nuclei, but they do not come to rest on it.

The placement of electrons in atoms determines the configuration of the periodic table, a complete version of which is printed at the end of this chapter. Elements are listed across this table in *periods* such that elements located in the same vertical *groups* have generally similar chemical behavior. The derivation of the complete periodic table showing more than 100 elements is too complicated for this book. So, in the remainder of this chapter, the first 20 elements will be discussed in order and the placement of electrons in the atoms of these elements will illustrate how these elements can be placed in the periodic table. From this information a brief 20-element periodic table will be constructed that should be very useful in explaining chemical behavior.

This page titled 2.5.3.1: Elements, Atoms, and Atomic Theory is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

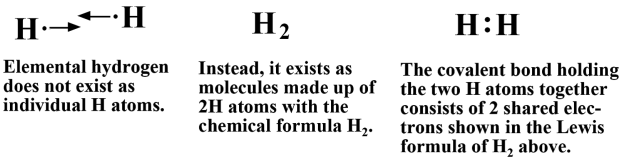
- **3.1: Elements, Atoms, and Atomic Theory** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.3.2: Hydrogen - The Simplest Atom

Hydrogen, **H**, is the element with atomic number 1. Most hydrogen atoms consist of a single proton forming the nucleus with 1 electron per hydrogen atom. Recall from Section 2.12 and Figure 2.6 that elemental hydrogen exists as molecules with 2 H atoms, chemical formula H2, in which the 2 H atoms are joined together by a *covalent bond* consisting of 2 shared electrons. Molecules consisting of 2 atoms so joined are called **diatomic** molecules. As will be seen, several important elements among the first 20 elements are gases that consist of diatomic molecules in their elemental forms.

Showing Electrons in Atomic Symbols and Molecular Formulas

In discussing chemical behavior related to atomic structure, it is particularly useful to have a means of showing the electrons in the atoms (more specifically, the less strongly held outer shell electrons). This is done with **Lewis symbols** (named after G. N. Lewis) also called **electron-dot symbols**. The Lewis symbol of the hydrogen at



H· Figure 3.1. Figure 3.1. Elemental hydrogen exists as molecules each of which contains two H atoms.

Properties and Uses of Elemental Hydrogen

Pure elemental H2 under normal conditions is a colorless, odorless gas that has the lowest density of any pure substance. Liquified H2 boils at a very cold -253˚C and solidifies at -259˚ C. Hydrogen gas is widely used in the chemical industry to react chemically with a large number of substances. It burns readily with a large release of energy and mixtures of hydrogen with oxygen or air are extremely explosive. The chemical reaction for elemental hydrogen burning with oxygen (O2) in air is

$$2\,H_2 + O_2 \longrightarrow 2\,H_2O + \text{energy} \qquad (2.5.3.2.1)$$

The product of this reaction is water. Used as a fuel, elemental hydrogen is a very green element because when it is burned or otherwise reacted to provide energy, the reaction product is simply water, H2O. Furthermore, given a source of electrical energy, elemental hydrogen and elemental oxygen can be produced at two separate electrodes by passing a direct current through water in which an appropriate salt has been dissolved to make the water electrically conducting:

$2H2O \xrightarrow[\text{current}]{\text{Electrical}} 2H2 + O2$ (3.2.2)

So elemental hydrogen generated by the application of electrical energy to water provides a source of energy that can be moved from one place to another and utilized to produce electricity in fuel cells (see below) or for other beneficial purposes such as the synthesis of ammonia essential as a source of plant fertilizer nitrogen. The production of elemental hydrogen by electrolysis may be regarded as a green process because it does not require any reagents other than water. Furthermore, the electrolysis byproduct oxygen is harmless and has many uses whereas hydrogen made by the reaction of steam with carbon-containing compounds (see below) consumes fossil fuel and generates CO, which in some cases is burned producing greenhouse gas carbon dioxide. The main disadvantage of the electrolysis process for H2 generation is the relatively low efficiency by which the electricity is used in the process and improvements are needed in this area.

Elemental hydrogen is widely used for chemical synthesis and other industrial applications. Its preparation by electrolysis of water was mentioned above. It is now most commonly prepared from methane, CH4, the main ingredient of natural gas, by **steam reforming** at high temperatures and pressures:

$CH4 + H2O \xrightarrow[30\text{ atm P}]{800°C\text{ T}} CO + 3H2$ (3.2.3)

Hydrogen is used to manufacture a number of chemicals. Two of the most abundantly produced chemicals that require hydrogen for synthesis are ammonia, NH_3, and methanol (methyl alcohol, CH_3OH). The latter is generated by the reaction between carbon monoxide and hydrogen:

$$CO + 2\,H_2 \longrightarrow CH_3OH \qquad (2.5.3.2.2)$$

Methanol used to be made by heating wood in the absence of air and condensing methanol from the vapor given off, a process known as destructive distillation. Generation of so-called wood alcohol made by this relatively green process from biomass has the potential to supply at least a fraction of the methanol now needed, thus reducing the consumption of natural gas.

Methanol has some important fuel uses. During the 1930s it was used instead of gasoline to run internal combustion engines to power a significant fraction of automobiles in France before Middle Eastern oil fields became such an abundant source of petroleum. At present it is blended with gasoline as an oxygenated additive; engines using this blended fuel produce less pollutant carbon monoxide. Now the most common use of methanol as a fuel is to break it down to elemental hydrogen and carbon dioxide to produce hydrogen used in fuel cells.

In addition to its uses in making ammonia and methanol, hydrogen is added chemically to hydrocarbon molecules in some fractions of gasoline to upgrade the fuel value of gasoline. Hydrogen can be added directly to coal or reacted with carbon monoxide to produce synthetic petroleum. It is also combined with unsaturated vegetable oils to make margarine and other hydrogenated fats and oils. This application is controversial and becoming less common because of suspected adverse long-term health effects of these products commonly called trans fats.

Hydrogen in Fuel Cells

Fuel cells, discussed further in Chapter 15, are devices that enable hydrogen to “burn” at around room temperature and to produce electricity directly without going through some sort of internal combustion engine and electricity generator. A fuel cell (Figure 3.2) consists of two electrically conducting electrodes, an anode and a cathode that are contacted with elemental H2 and O2, respectively. As shown in the diagram, at the anode H2 loses electrons (it is said to be oxidized) to produce H+ ion. At the cathode O2 gains electrons (it is said to be reduced) and reacts with H+ ions to produce water, H2O. The H+ ions required for the reaction at the cathode are those generated at the anode and they migrate to the cathode through a solid membrane permeable to protons (the H+ ion is a proton). The net reaction is

$$2\,H_2 + O_2 \longrightarrow 2\,H_2O + \text{electrical energy} \qquad (2.5.3.2.3)$$

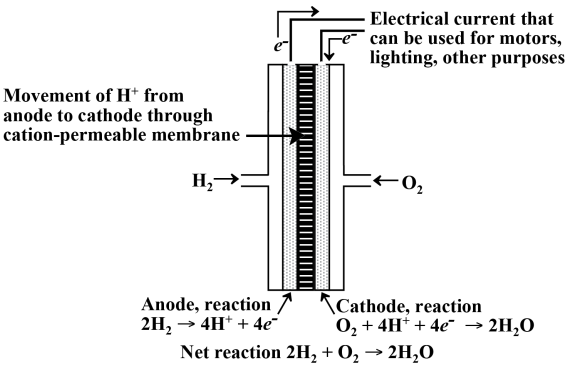


Figure 3.2. Cross-sectional diagram of a fuel cell in which elemental hydrogen can be reacted with elemental oxygen to produce electricity directly with water as the only chemical product.

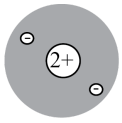
This page titled 2.5.3.2: Hydrogen - The Simplest Atom is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **3.2: Hydrogen - The Simplest Atom** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.3.3: Helium - The First Noble Gas

The second element in the periodic table is helium, He, atomic number 2. All helium atoms have 2 protons in their nuclei and 2 electrons. There are two isotopes of helium of which $^{4}_{2}He$ that contains 2 neutrons in the nucleus is by far the predominant one, with much smaller numbers of the lighter isotope, $^{3}_{2}He$ which has 2 protons and 1 neutron in its nucleus and a mass number of 3.

Helium is a *noble gas* meaning that it exists only as atoms of the elements that are never bonded to other atoms. Figure 3.3 is a representation of the helium atom showing its 2 electrons. The Lewis symbol of helium is simply He with 2 dots. This shows a very important characteristic of atoms. As electrons are added to atoms with increasing atomic number, they are added at various levels known as **electron shells**. The one electron in hydrogen, H, goes into the first electron shell, the one with the lowest possible energy. The second electron added to make the helium atom also goes into the first electron shell. This lowest electron shell can contain a maximum of only 2 electrons, so helium has a **filled electron shell**. Atoms with filled electron shells have no tendency to lose, gain, or share electrons and, therefore, do not become involved with other atoms through chemical bonding. Such atoms exist alone in the gas phase and the elements of which they consist are called noble gases. Helium is the first of the noble gases.



A helium atom has a filled electron shell containing 2 electrons.



It can be represented by the Lewis symbol above.

Figure 3.3. Two representations of the helium atom having a filled electron shell.

Helium gas has a very low density of only 0.164 g/L at 25˚C and 1 atm pressure. Elemental helium is the second least dense substance next to hydrogen gas. It is this low density characteristic that makes helium so useful in balloons, including weather balloons, which can stay aloft for days, reaching very high altitudes.

Helium is pumped from the ground with some sources of natural gas, some of which contain up to 10% helium by volume. Helium was first observed in the sun by the specific wavelengths of light emitted by hot helium atoms. Underground sources of helium were discovered by drillers searching for natural gas in southwestern Kansas who tried to ignite gas from a new well and were disappointed to find that it would not burn, since it was virtually pure helium!

Chemically unreactive, helium has no chemical uses, except to provide a chemically inert atmosphere. A nontoxic, odorless, tasteless, colorless gas, helium is used because of its unique physical properties. Applications in weather balloons and airships were mentioned previously. Because of its low solubility in blood, helium is mixed with oxygen for breathing by deep-sea divers and persons with some respiratory ailments. Use of helium by divers avoids the very painful condition called “the bends” caused by bubbles of nitrogen forming from nitrogen gas dissolved in blood.

The greatest use of helium is as the super-cold liquid, which boils at a temperature of only 4.2 K above absolute zero (-269˚C), especially in the growing science of **cryogenics**, which deals with very low temperatures. Some metals are superconductors at such temperatures so that helium is used to cool electromagnets enabling relatively small magnets to develop very powerful magnetic fields. Such magnets are components of the very useful chemical tool known as nuclear magnetic resonance (NMR). The same kind of instrument modified for clinical applications and called MRI is used as a medical diagnostic tool for scanning sections of the body for evidence of tumors and other maladies.

Hydrogen Wants to be Like Helium

Examination of the Lewis symbol of helium (right, Figure 3.3) and the Lewis formula of elemental hydrogen, H2, (Figure 3.1) shows that each of the two hydrogen atoms in the H2 molecule can lay claim to 2 electrons and thereby come to resemble the helium atom. Recall that helium is a noble gas that is very content with its 2 electrons. Each of the H atoms in H2 is satisfied with 2 electrons, though they are shared. This indicates a basic rule of chemical bonding that atoms of an element tend to acquire the same electron configuration as that of the nearest noble gas. In this case, hydrogen, which comes just before helium in the periodic table, gains the noble gas configuration of helium by sharing electrons.

This page titled 2.5.3.3: Helium - The First Noble Gas is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **3.3: Helium - The First Noble Gas** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

https://bio.libretexts.org/@go/page/108776

2.5.3.4: Lithium, The First Metal

The element with atomic number 3 is lithium (Li), atomic mass 6.941. The most abundant lithium isotope is $^{7}_{3}Li$ having 4 neutrons in its nucleus. A few percent of helium atoms are the $^{6}_{3}Li$ isotope, which has only 3 neutrons. The third electron in lithium cannot fit in the lowest energy shell, which, as noted above, is full with only 2 electrons. Therefore, the third electron in lithium goes into a second shell, that is, an **outer shell**.

As a consequence of its electronic structure, lithium is the lowest atomic number element that is a **metal**. In a general sense, metals are elements that normally have only 1–3 electrons in their outer shells. These electrons can be lost from metals to produce positively charged *cations* with charges of +1, +2, or +3. In the pure elemental state metals often have a characteristic **luster** (shine), they are **malleable** (can be flattened or pushed into various shapes without breaking) and they **conduct electricity**. Although some metals, notably lead and mercury, are very dense, lithium is the least dense metal at only 0.531 g/cm3.

Two of lithium's 3 electrons are **inner electrons** contained in an **inner shell** as in the immediately preceding noble gas helium. Inner shell electrons such as these stay on average relatively close to the nucleus, are very tightly held, and are not exchanged or shared in chemical bonds. As mentioned above, the third electron in lithium is an **outer electron** farther from, and less strongly attracted to, the nucleus. The outer electron is said to be in the atom's *outer shell*. These concepts are illustrated in Figure 3.4.

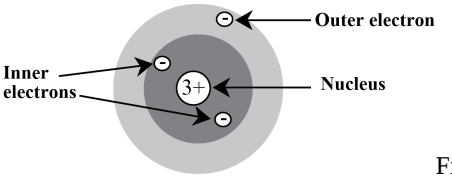


Figure 3.4. An atom of lithium, Li, has 2 inner electrons and 1 outer electron. The latter can be lost to another atom to produce the Li^+ ion, which is present in ionic compounds (see Section 2.12).

The Lewis symbol for atoms such as lithium that have both inner shell and outer shell electrons normally shows just the latter. (Inner shell electrons can be shown on symbols to illustrate a point, but normally this takes too much space and can be confusing.) Since lithium has only one outer shell electron, its Lewis symbol is

$$\mathbf{Li\cdot}$$

Consider that the lithium atom has an inner shell of 2 electrons, just like helium. Being only 1 electron away from the helium noble gas structure, lithium has a tendency to lose its extra electron so it can be like helium as shown by the following:

$$\mathbf{Li(:)\cdot \rightarrow Li:^{+} + e^{-}} \quad (3.4.1)$$

Note that the product of this reaction is no longer a neutral atom, but is a positively charged Li^+ *cation*. In losing an electron to become a cation, the lithium atom is said to be **oxidized**. When lithium forms chemical compounds with other elements, it does so by losing an electron from each lithium atom to become Li^+ cations. These, then, are attracted to negatively charged *anions* in ionic compounds.

Lithium compounds have a variety of uses. Lithium carbonate, Li2CO3, is widely prescribed as a pharmaceutical to alleviate the symptoms of mania in manic-depressive and schizo-affective mental disorders. Lithium carbonate is the most common starting material for the preparation of other lithium compounds and is an ingredient of specialty glasses and enamels and of ceramic ware that expands only minimally when heated. Lithium hydroxide, LiOH, is used to formulate some kinds of lubricant greases. In combination with iodine, lithium has been used to make cells that are sources of electricity for cardiac pacemakers. Implanted in the patient's chest, some of these pacemakers and their batteries have lasted for 10 years before having to be replaced.

Long an element with limited uses, lithium has become a rather "exciting" metal in the newly emerging sustainability economy because lithium-based storage batteries in which Li^+ ion is a charge carrier have become the storage batteries of choice for computers, portable electric devices and, especially, electric and hybrid automobiles. Lithium storage batteries exhibit superior qualities with respect to charge held per unit mass, stability, and longevity. In addition, lithium dry cells in which Li metal is irreversibly converted to Li^+ ion during discharge have become attractive (though expensive) options in the throwaway dry cell market because they are very long lived and carry much more charge per unit mass than standard alkaline dry cells.

The most common source of lithium is lithium brines where lithium has become concentrated by leaching of rock and evaporation of water from highland salt flats in South America and western China. Lithium salts are collected for processing from the evaporation of water from the brines in evaporation ponds. Bolivia is the largest producer of lithium and there is significant potential of production from Chile, Argentina, Australia, China, and perhaps even the state of Nevada. As use of lithium in batteries increases, recycled lithium from spent batteries will become an important source of the element.

This page titled 2.5.3.4: Lithium, The First Metal is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **3.4: Lithium, The First Metal** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.3.5: The Second Period of the Periodic Table

The first period of the periodic table is a short one consisting of only two elements, hydrogen and helium. Lithium, atomic number 3 begins the second period, which has 8 elements. Elements with atomic numbers 4-10, which complete this period, are discussed in this section.

Beryllium, Atomic Number 4

Like atoms of all the elements in the second period of the periodic table, **beryllium**, atomic number 4, atomic mass 9.012, has 2 inner shell electrons. Beryllium also has 2 outer shell electrons, so its Lewis symbol is



In addition to 4 protons in their nuclei, beryllium atoms also have 5 neutrons. When the beryllium atom is oxidized to form a beryllium cation, the reaction is

Be: → Be^{2+} + 2e- (lost to another atom) (3.5.1)

Since the beryllium atom needs to lose 2 electrons to reach the two-electron helium electron configuration, it produces doubly charged Be^{2+} cations.

Beryllium has some important uses in metallurgy. Melted together with other metals, a process that produces **alloys**, beryllium yields metal products that are hard and corrosion-resistant. Beryllium alloys can be blended that are good electrical conductors and that are nonsparking when struck, an important characteristic in applications around flammable vapors. Among the devices for which beryllium alloys are especially useful are various specialty springs, switches, and small electrical contacts. Beryllium has found widespread application in aircraft brake components where its very high melting temperature (about 1290˚C) and good heat absorption and conduction properties are very advantageous.

In a sense, beryllium is somewhat the opposite of a green element. This is because of its adverse health effects, including **berylliosis**, a disease marked by lung deterioration. Because of the extreme inhalation hazard of Be, allowable atmospheric levels are very low. Many workers were occupationally exposed to beryllium as part of the nuclear reactor and weapons industry in the U.S. in the decades following World War II. In recognition of the adverse health effects of occupational exposure to beryllium, in the late 1990s the U.S. Government agreed to compensate workers suffering occupational exposure to this metal.

Boron, a Metalloid

Boron, B, atomic number of 5, atomic mass 10.81, consists primarily of the isotope with 6 neutrons in addition to 5 protons in its nucleus; a less common isotope has 5 neutrons. Two of boron's 5 electrons are in a helium core and 3 are outer electrons as denoted by



Boron is the first example of an element with properties intermediate between those of metals and nonmetals called **metalloids**. In addition to boron, the metalloids include silicon, germanium, arsenic, antimony, and tellurium, of which the most notable, silicon, is among the first 20 elements. In the elemental state, metalloids have a luster like metals, but they do not readily form simple cations. Unlike metals, which generally conduct electricity well, metalloids usually conduct electricity poorly, if at all, but can become conductors under certain conditions. Such materials are called **semiconductors** and are of crucial importance because they form the basis of the world's vast semiconductor industry, which has given us small, powerful computers and a huge array of other electronic products.

Boron is a high-melting substance (2190˚C) that is alloyed with copper, aluminum, and steel metals to improve their properties. As a good absorber of neutrons, boron is used in nuclear reactor control rods and as a neutron-absorbing additive to the water that circulates through a reactor core as a heat transfer medium. Boron nitride, BN, is extraordinarily hard, as are some other compounds of boron. Boron oxide, B_2O_3, is an ingredient of heat-insulating fiber glass and boric acid, H_3BO_3, is used as a flame retardant in cellulose insulation in houses.

The Element of Life, Carbon

Carbon, C, atomic number 6, brings us to the middle of the second period of the periodic table. In addition to its 2 inner electrons, the neutral carbon atom has 4 outer electrons as shown by the Lewis symbol



Figure 3.5 shows three organic compounds composed only of carbon and hydrogen (hydrocarbons), each containing 8 carbon atoms. These structures illustrate the bonding diversity of carbon.

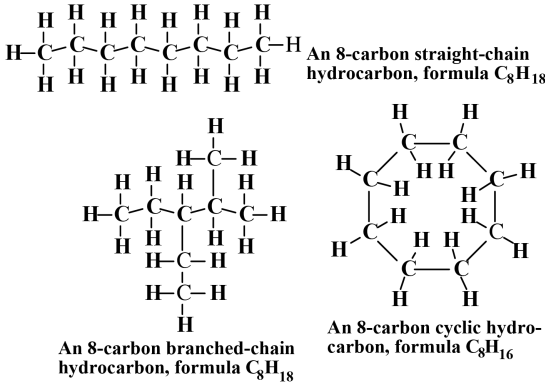


Figure 3.5. Illustration of the bonding versatility of carbon with three carbon-containing hydrocarbon compounds arranged as straight (continuous) chains, branched chains, and in a cyclic compound.

The ability of carbon atoms to bond with each other determines the properties of the several important and useful forms of elemental carbon. (Different forms of the same element are called allotropes.) Very fine carbon powder composes **carbon black**, which is used in tires, inks, and printer toner. Carbon atoms bonded in large flat molecules compose **graphite**, so soft and slick that it is used as a lubricant. Carbon treated with steam or carbon dioxide at elevated temperatures develops pores that give the carbon an enormous surface area. This product is **activated carbon** that is very useful in purifying foods, removing organic pollutants from water and removing pollutant vapors from air. Elemental carbon fibers are bonded together with epoxy resins to produce light composites so strong that they are used for aircraft construction. Bonded together in a different way that gives a very hard and rigid structure, carbon atoms produce diamond.

A particularly interesting class of carbon allotropes is that of **fullerenes** consisting of elemental carbon bonded in generally 5- and 6-membered rings to form spheres, ellipsoids, and tubes. The first of this class of elemental carbon discovered only in 1985 consists of aggregates of 60 carbon atoms bonded together in 5- and 6-membered rings that compose the surface of a sphere. This structure resembles the geodesic domes designed as building structures by Buckminster Fuller, a visionary designer. Therefore, the discoverers of this form of carbon named it buckminsterfullerene and the C60 balls, which resemble soccer balls in their structure, are commonly called “buckyballs.” Since the discovery of the C60 fullerene, many related forms have been synthesized, of which the most interesting may be very narrow carbon tubes called **carbon nanotubes**. (“Nano” is a prefix commonly assigned to materials in which the individual units have dimensions around 1 nanometer or 1 × 10-9 meters.) Carbon nanotubes have very interesting properties including some forms with an extraordinarily great length-to-diameter ratio of up to 132,000,000:1. Because of their unique dimensions, extraordinary strength, electrical properties, and ability to efficiently conduct heat, carbon nanotubes are of intense interest in materials science, nanotechnology, electronics, optics, and other high technology applications. However, special attention needs to be given to their potential toxicity.

Green Carbon from the Air

Carbon is present in the air as gaseous carbon dioxide, CO2. Although air is only about 0.04% CO2 by volume, it serves as the source of carbon for the growth of green plants. In so doing, the chlorophyll in plants captures solar energy in the form of visible light, represented *hν*, and uses it to convert atmospheric carbon dioxide to high-energy glucose sugar, C6H12O6, as shown by the following reaction:

$$6\,CO_2 + 6\,H_2O \xrightarrow{h\nu} C_6H_{12}O_6 + 6\,O_2 \qquad (2.5.3.5.1)$$

The carbon fixed in the form of C6H12O6 and related compounds provides the basis of the food chains that sustain all organisms. Organic carbon produced by photosynthesis in eons past also provided the raw material for the formation of

petroleum, coal, and other fossil fuels. Now, as supplies of these scarce resources dwindle and as the environmental costs of their extraction, transport, and utilization mount, there is much renewed interest in photosynthetically produced carbon compounds as raw materials and even fuels. Despite the low levels of carbon dioxide in the atmosphere and the relatively low efficiency of photosynthesis, rapidly growing plants, such as some varieties of hybrid poplar trees, can produce enormous quantities of carbon compounds very rapidly and in a sustainable manner.

Nitrogen from the Air

Nitrogen, N, atomic number 7, atomic mass 14.01, composes 78% by volume of air in the form of diatomic N2 molecules. The nitrogen atom has 7 electrons, 2 contained in its inner shell and 5 in its outer shell. So its Lewis symbol is the following:



Figure 3.6.

Nitrogen gas does not burn and is generally chemically unreactive. Advantage is taken of the extreme chemical stability of nitrogen gas in applications where a nonreactive gas is needed to prevent fires and explosions. Although almost 80% of the air that people breathe consists of elemental nitrogen gas, people have died of asphyxiation by entering areas filled with nitrogen gas in which oxygen is absent. Since nitrogen gas has no odor, it does not warn of its presence.

Huge quantities of liquid nitrogen, which boils at a very cold -190˚C, are used in areas where cold temperatures are needed. This frigid liquid is employed to quick-freeze foods and for drying materials in freeze-drying processes, Biological materials, such as semen used in artificial breeding of animals, can be preserved in liquid nitrogen.

The atmosphere is an inexhaustible reservoir of nitrogen. However, it is very difficult to get nitrogen into the chemically combined form in which it occurs in simple inorganic compounds or proteins. This is because of the extreme stability of the N2 molecule, mentioned above. The large-scale chemical fixation of atmospheric nitrogen over a catalytic surface at high temperatures and pressure as represented by the reaction

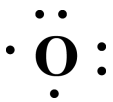
$$\mathrm{N_2 + 3\,H_2 \longrightarrow 2\,NH_3} \qquad (2.5.3.5.2)$$

was a major accomplishment of the chemical industry about a century ago. It enabled the large-scale production of relatively cheap nitrogen fertilizers that resulted in highly increased crop production, as well as the manufacture of enormous quantities of nitrogen-based explosives that made possible the unprecedented carnage of World War I. Despite the extreme conditions required for the preparation of nitrogen compounds by humans in the anthrosphere, humble bacteria accomplish the same thing under ambient conditions of temperature and pressure, converting N2 from the air into organically bound nitrogen in biomass. Prominent among the bacteria that do this are *Rhizobium* bacteria that grow symbiotically on the roots of legume plants, fixing atmospheric nitrogen that the plants need and drawing nutrients from the plants. Because of this ability, legumes, such as soybeans and clover grow well with less artificial nitrogen fertilizer than that required by other plants. One of the exciting possibilities with genetically modified plants is the potential to develop nitrogen-fixing varieties of corn, wheat, rice, and other crops that now lack the capability to fix nitrogen.

Nitrogen is an essential life element that is present in all proteins, hemoglobin, chlorophyll, enzymes, and other life molecules. It circulates through nature in the **nitrogen cycle** by which elemental nitrogen is incorporated from the atmosphere into biological material. Nitrogen-containing biomass is converted during biodegradation by bacteria to inorganic forms, which may be utilized as nutrient nitrogen by plants. Eventually, bacterial processes convert the nitrogen back to elemental N2, which is returned to the atmosphere to complete the cycle.

Oxygen, the Breath of Life

Oxygen, atomic number 8, atomic mass 16.00 is required by humans and many other living organisms. A diatomic nonmetal, elemental oxygen consists of O2 molecules and makes up 21% of the volume of air. Of its 8 electrons, the oxygen atom has 6 in the outer shell as represented by the Lewis formula:



Oxygen can certainly be classified as a green element for a number of reasons, not the least of which is that O in the atmosphere is there for the taking. Elemental oxygen is transferred from the atmosphere to the anthrosphere by liquifying air and distilling the liquid air, the same process that enables isolation of pure nitrogen. These two gases are also separated from air by adsorption onto solid surfaces under pressure followed by removal under vacuum. Pure oxygen has a number of applications including use as a gas to breathe by people with lung deficiencies, in chemical synthesis, and in oxyacetylene torches employed for welding and cutting metals.

Although the elemental oxygen molecule is rather stable, at altitudes of many kilometers in the stratosphere, it is broken down to oxygen atoms by the absorption of ultraviolet radiation from the sun as shown in Chapter 2, Reaction 2.13.1 As illustrated by Reaction 2.13.2, the oxygen atoms formed by the photochemical dissociation of O2 combine with O2 molecules to produce molecules of ozone, O3. The result is a layer of highly rarefied air containing some ozone over an altitude range of many kilometers located high in the stratosphere. There is not really much ozone in this layer. If it were pure ozone under the conditions of pressure and temperature that occur at ground level, the ozone layer would be only about 3 millimeters thick! This stratospheric ozone, sparse though it is, serves an essential function in protecting organisms on Earth's surface from the devastating effects of ultraviolet radiation from the sun. Were it not for stratospheric ozone, life as it is now known could not exist on Earth.

Ozone has a split personality as a green form of oxygen. As discussed above, ozone in the stratosphere is clearly beneficial and essential for life. But it is toxic to inhale at levels less than even one part per million by volume. Ozone is probably the most harmful constituent of air polluted by the formation of photochemical smog in the atmosphere at ground levels.

The most notable chemical characteristic of oxygen is its ability to combine with other materials in energy-yielding reactions. One such reaction with which most people are familiar is the burning of gasoline in an automobile,

$$2\,C_8H_{18} + 25\,O_2 \longrightarrow 16\,CO_2 + 18\,H_2O + \text{energy} \qquad (2.5.3.5.3)$$

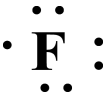
performed so efficiently that the combustion of only 1 gallon of gasoline can propel a full-size automobile more than 25 miles at highway speeds. Along with many other organisms, we use oxygen in our bodies to react with nutrient carbohydrates,

$$C_6H_{12}O_6\ (\text{glucose}) + 6\,O_2 \longrightarrow 6\,CO_2 + 6\,H_2O + \text{energy} \qquad (2.5.3.5.4)$$

to provide energy that we use. Whereas the combustion of a fuel such as gasoline occurs at red-hot temperatures, the "burning" of carbohydrates in our body occurs through the action of enzymes in the body at body temperature of only 37˚C.

The Most Nonmetallic Element, Fluorine

Fluorine, F, atomic number 9, atomic mass 19.00 has 7 outer electrons as shown by its Lewis symbol



Elemental fluorine consists of diatomic F2 molecules constituting a greenish-yellow gas. Fluorine is the most nonmetallic of all the elements. It reacts violently with metals, organic matter, and even glass! Elemental fluorine is a very corrosive poison that attacks flesh and forms wounds that heal very poorly. Because of its hazards, the practice of green chemistry seeks to minimize the generation or use of F2.

Fluorine is used in chemical synthesis. It was once widely employed to make Freons, chlorofluorocarbon compounds such as dichlorodifluoromethane, Cl2CF2, that were used as refrigerant fluids, spray can propellants, and plastic foam blowing agents. As discussed in Chapter 10, these compounds were found to be a threat to the vital stratospheric ozone layer mentioned in the discussion of oxygen above. They have now been replaced with fluorine-containing substitutes such as HFC-134a, CH2FCF3, which either do not contain the chlorine (Cl) that destroys stratospheric ozone or undergo destruction by atmospheric chemical processes near Earth's surface, and thus never reach the stratosphere.

This page titled 2.5.3.5: The Second Period of the Periodic Table is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **3.5: The Second Period of the Periodic Table** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons

Only one element remains to complete the second period of the abbreviated periodic table. This element is **neon**, atomic number 10, atomic mass 20.18. Although most neon atoms have 10 neutrons in addition to the 10 protons in their nuclei, some have 12 neutrons, and very few have 11. Neon is a gas consisting of individual Ne atoms that constitutes about 2 parts per thousand of the volume of air. Neon is recovered by distillation of liquid air. Its most common use is as a gas in tubes through which an electrical discharge is carried in glowing neon signs.

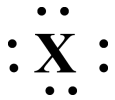
The total of 10 electrons in the neon atom are contained in two shells with 8 in the outer shell. So the Lewis symbol of neon is



With 8 electrons, the second shell of electrons is a *filled electron shell*. (Recall that helium has a filled electron shell with only 2 electrons.) Because it has a filled outer electron shell, neon is a noble gas that exists as single gas-phase atoms and forms no chemical compounds.

The Special Significance of the Octet

In addition to helium and neon, there are four other noble gas elements. These are argon (atomic number 18), krypton (atomic number 36), xenon (atomic number 54), and radon (atomic number 86). Other than helium, these all share a common characteristic of 8 outer-shell electrons. Such an electron configuration can be shown by the general Lewis symbol,



and is known as an **octet** of electrons, where X is the chemical symbol of the noble gas.

Although only atoms of noble gases have octets as single atoms, many other atoms acquire octets by forming chemical bonds. This tendency of atoms to acquire stable octets through chemical bonding is the basis of the **octet rule**, which is used in this book to predict and explain the chemical bonding in a number of compounds, such as those discussed in Chapter 4. To see a simple application of the octet rule in chemical bonding, consider the bonds involved in molecular elemental nitrogen, N2. Recall the Lewis symbol of the N atoms showing 5 dots to represent the 5 outer-shell electrons in each N atom. Figure 3.6 shows bond formation between 2 N atoms.

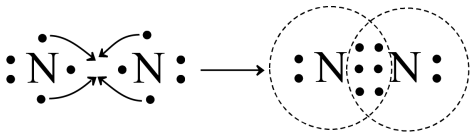


Figure 3.6. Illustration of the octet rule in the covalent bonding together of the 2 N atoms in N2. The six electrons in the bond connecting the two N atoms constitute a triple bond.

Each of the two N atoms in the N2 molecule needs 8 outer-shell electrons, but only 10 electrons are available to provide these electrons (the two inner-shell electrons in each N atom are not available to form bonds). This means that a lot of the electrons will have to be shared to form a bond between the N atoms in the N2 molecule, and, in fact, 6 of the 10 outer-shell electrons available in two N atoms are shared between the N atoms to give a *triple bond* consisting of 3 pairs of shared electrons as shown by the 3 pairs of dots between the 2 N atoms in the molecule of N2 in Figure 3.6.

This page titled 2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **3.6: The Magic Octet of 8 Outer-Shell Electrons** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.3.7: Completing the 20-Element Periodic Table

So far in this chapter 10 elements have been defined and discussed. A total of 10 more are required to complete the 20-element abbreviated periodic table. Their names and properties are summarized briefly here. The periodic table is given in Figure 3.9.

Sodium, **Na**, atomic number 11, atomic mass 22.99, comes directly below lithium in the periodic table and is very similar to lithium in being a soft, chemically very reactive metal. There is one major isotope of sodium containing 12 neutrons in the atom's nucleus. Sodium has 10 inner-shell electrons contained in its first inner shell of 2 electrons and its second one of 8 electrons. The 11th electron in the sodium atom is in a third shell, which is an outer shell. This is shown as a single dot in the Lewis symbol of Na in Figure 3.7. The electrons in sodium can be represented as shown in Figure 3.7.

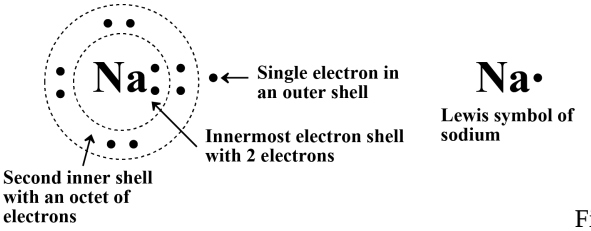


Figure 3.7. Representation of the electrons in 2 inner shells and 1 outer shell of sodium and the Lewis symbol of sodium.

Magnesium, **Mg**, atomic number 12, atomic mass 24.31, has 12 electrons per atom so it has 2 outer shell electrons. There are three isotopes of magnesium containing 12, 13, and 14 neutrons. Magnesium is a relatively strong, very lightweight metal that is used in aircraft, extension ladders, portable tools, and other applications where light weight is particularly important.

Aluminum, **Al**, atomic number 13, atomic mass 26.98, has 3 outer-shell electrons in addition to its 10 inner electrons. Aluminum is a lightweight metal used in aircraft, automobiles, electrical transmission lines, building construction and many other applications. Although it is chemically reactive, the oxide coating formed when aluminum on the surface of the metal reacts with oxygen in air is self-protecting and prevents more corrosion.

In some important respects aluminum can be regarded as a green metal. This is because aluminum enables construction of strong lightweight components which, when used in aircraft and automobiles, require relatively less energy to move. So aluminum is important in energy conservation. Aluminum cables also provide an efficient way to transmit electricity. Although the ores from which aluminum is made are an extractive resource dug from the earth, aluminum is an abundant element. And there are alternative resources that can be developed, including aluminum in the fly ash left over from coal combustion. Furthermore, aluminum is one of the most recyclable metals, and scrap aluminum is readily melted down and cast into new aluminum goods.

If an "element of the century" were to be named for the 1900s, humble **silicon**, **Si**, atomic number 14, atomic mass 28.09, would be a likely candidate. This is because silicon is the most commonly used of the semiconductor elements and during the latter 1900s provided the basis for the explosion in electronics and computers based upon semiconductor devices composed primarily of silicon. Despite the value of these silicon-based products, silicon is abundant in soil and rocks, ranking second behind oxygen as a constituent of Earth's crust. The silicon atom has 4 outer-shell electrons, half an octet, and it is a metalloid, intermediate in behavior between the metals on the left of the periodic table and the nonmetals on the right.

By vastly reducing the bulk of electronic components relative to performance, silicon has contributed to a huge saving of materials used in radios, televisions, communications equipment, and other electronic devices. Furthermore, the silicon-based semiconductor devices used in solid-state electronics consumes only a fraction of the electricity once used by vacuum tube based devices. The bulky wires made of relatively scarce copper formerly employed for transmitting communications signals electrically have been largely replaced by fiber optic devices consisting of transparent silica, SiO2, which transfer information as pulses of light. A hair-like optical fiber can transmit many times the amount of information per unit time as the thick copper wire that it replaces. And the energy required for transmission of a unit of information by a fiber optic cable is minuscule compared to that required to send the same information by electrical impulse over copper wire. So silicon is truly a green element that, although cheap and abundant, performs electronic and communications functions much faster and better than the copper and other metals that it has replaced.

Phosphorus, **P**, atomic number 15, atomic mass 30.97, has 5 outer-shell electrons. So it is directly below nitrogen in the periodic table and resembles nitrogen in its chemical behavior. Pure elemental phosphorus occurs in several forms, the most abundant of

which is white phosphorus. White phosphorus is a chemically very reactive nonmetal that may catch fire spontaneously in the atmosphere. It is toxic and causes deterioration of bone. The jawbone is especially susceptible to the effects of phosphorus and develops a condition known as “phossy jaw” in which the bone becomes porous and weak and may break from the strain of chewing. Chemically combined phosphorus is an essential life element, however, and is one of the components of DNA, the basic molecule that directs molecular life processes. Phosphorus is also an essential plant fertilizer and is an ingredient of many industrial chemicals including some pesticides.

Arsenic is in the same group of the periodic table as phosphorus and occurs as an impurity with phosphorus processed from ore. If this phosphorus is to be used for food, the arsenic has to be removed.

Sulfur, **S**, atomic number 16, atomic mass 32.06, has 6 outer-shell electrons. It is a brittle, generally yellow nonmetal. It is an essential nutrient for plants and animals occurring in the amino acids that compose proteins. Sulfur is a common air pollutant emitted as sulfur dioxide, SO2, in the combustion of fossil fuels that contain sulfur. Much of the large quantities of sulfur required for industrial production of sulfuric acid and other sulfur-containing chemicals is reclaimed from the hydrogen sulfide, H2S, that contaminates much of the natural gas (methane, CH4) that is such an important fuel and raw material in the world today. In keeping with the best practice of green chemistry, the hydrogen sulfide is separated from the raw natural gas and about 1/3 of it is burned,

$$2\,H_2S + 3\,O_2 \longrightarrow 2\,SO_2 + 2\,H_2O \qquad (2.5.3.7.1)$$

producing sulfur dioxide, SO2. The sulfur dioxide produced is then reacted with the remaining hydrogen sulfide through the Claus reaction, below, yielding an elemental sulfur product that is used to synthesize sulfuric acid and other sulfur chemicals.

$$2\,H_2S + SO_2 \longrightarrow 3\,S + 2\,H_2O \qquad (2.5.3.7.2)$$

Chlorine, **Cl**, atomic number 17, atomic mass 35.453, has 7 outer-shell electrons, just 1 electron short of a full octet. Elemental chlorine is a greenish-yellow diatomic gas consisting of Cl2 molecules. In these molecules the Cl atoms attain stable octets of outer-shell electrons by sharing two electrons in a covalent bond as illustrated in Figure 3.8. The chlorine atom can also accept an electron to attain a stable octet in the Cl- anion as shown in the ionic compound sodium chloride, NaCl, in Figure 3.8.

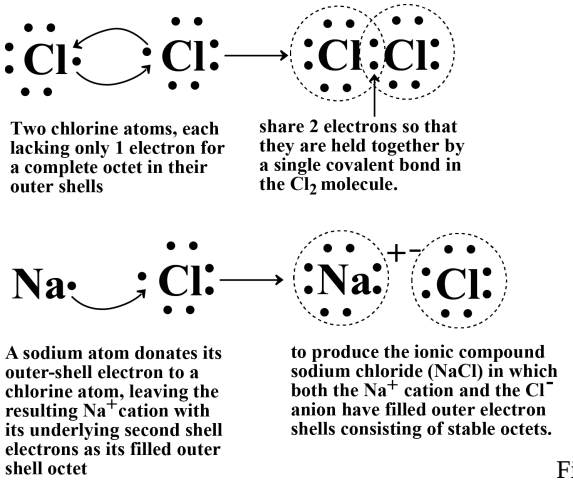


Figure 3.8. Chlorine atoms can attain a stable octet of outer-shell electrons by sharing electrons in covalent bonds, as shown for molecular diatomic elemental Cl2, or by accepting electrons to produce Cl- anion as shown for ionic sodium chloride, NaCl.

Elemental chlorine can be deadly when inhaled and was the first military poison gas used in World War I. Despite its toxic nature, chlorine gas has saved many lives because of its use for approximately the last 100 years as a drinking water disinfectant that has eradicated deadly water-borne diseases, such as cholera and typhoid. Chlorine is an important industrial chemical that is used to make plastics and solvents. There is no possibility of a shortage of chlorine and it can even be made by passing an electrical current through seawater, which contains chlorine as dissolved sodium chloride.

The green aspects of chlorine depend upon its application. Elemental chlorine is certainly a dangerous material whose manufacture and use are generally to be avoided where possible in the practice of green chemistry. But, as noted above, elemental chlorine has saved many lives because of its uses to disinfect water. A number of persistent pesticides including DDT are organic compounds composed of chlorine along with carbon and hydrogen. In addition to the ecological damage done by these pesticides, the waste

byproducts of their manufacture and of the production of other organochlorine compounds are among the most abundant contaminants of troublesome hazardous waste chemical dumps. A common plastic, polyvinyl chloride (PVC), contains chlorine. This plastic is widely used in water pipe and drain pipe, in the former application replacing relatively scarce and expensive copper metal and toxic lead. But the material used to make PVC is volatile vinyl chloride. It is one of the few known human carcinogens, having caused documented cases of a rare form of liver cancer in workers formerly exposed to very high levels of vinyl chloride vapor in the workplace. Because of the dangers of elemental chlorine and the problems caused by organochlorine compounds, the practice of green chemistry certainly tries to minimize the production and use of elemental chlorine and generally attempts to minimize production of organochlorine compounds and their dispersion in the environment.

Element number 18, **argon**, **Ar**, atomic mass 39.95, brings us to the end of the third period of the abbreviated periodic table. It has a complete octet of outer-shell electrons and is a noble gas. No true chemical compounds of argon have been isolated and no chemical bonds involving this element were known until formation of a very unstable transient bond involving Ar atoms was reported in September, 2000. Argon composes about 1% by volume in the atmosphere. Largely because of its chemically inert nature, argon has some uses. It is employed as a gas to fill incandescent light bulbs. In this respect it helps prevent evaporation of white-hot tungsten atoms from the glowing lamp filament, thus significantly extending bulb life. It is also used as a plasma medium in instruments employed for inductively coupled plasma atomic emission analysis of elemental pollutants. In this application a radiofrequency signal is used to convert the argon to a gaseous plasma that contains positively charged Ar+ ions and negatively charged electrons and is heated to an incredibly hot 10,000˚C.

Completing the Periodic Table

The next element to be added to the abbreviated periodic table is element number 19. This begins a fourth period of the periodic table. This period actually contains 18 elements, but we will take it only as far as the first two. That is because element number 21 is the first of the transition metals and to explain their placement in the periodic table on the basis of the electrons in them gets a little more complicated and involved than is appropriate for this book. The reader needing more details is referred to other standard books on beginning chemistry.2,3

The element with atomic number 19 is **potassium**, **K**, having an atomic mass of 39.10. Most potassium consists of the isotope with 20 neutrons, $^{39}_{19}K$. However, a small fraction of naturally occurring K is in the form of $^{40}_{19}K$. This is a radioactive isotope of potassium and since we all have potassium (an essential element for life) in our bodies, we all are naturally radioactive! Muscle mass contains more potassium than adipose (fat) tissue, so more muscular people are more radioactive. But not to worry, the levels of radioactivity from potassium in the body are too low to cause concern and, under any circumstances, cannot be avoided. (One proponent of nuclear energy has pointed out that sleeping with a muscular person exposes one to more radioactivity than does living close to a nuclear power reactor.)

The same things that can be said of sodium, element number 11, are generally true of potassium. In the pure elemental state, potassium is a very reactive *alkali metal*. As an essential element for life, it is a common fertilizer added to soil to make crops grow well. Chemically, potassium loses its single outer-shell electron to produce K+ ion.

Calcium, **Ca**, atomic number 20, atomic mass 40.08, has 2 outer-shell electrons, two beyond a full octet. The calcium atom readily loses its 2 “extra” electrons to produce Ca2+ cation. Like other elements in its group in the periodic table, calcium is an alkaline earth metal. Elemental calcium metal is chemically reactive, though not so much so as potassium. Calcium has chemical properties very similar to those of magnesium, the alkaline earth metal directly above calcium in the periodic table.

Calcium is essential for life, although most soils contain sufficient calcium to support optimum crop growth. Calcium is very important in our own bodies because as hard mineral hydroxyapatite, Ca5OH(PO4)3, it is the hard material in teeth and bones. Calcium deficiency can cause formation of poor teeth and the development of disabling osteoporosis a condition characterized by weak bones that is especially likely to afflict older women.

This page titled 2.5.3.7: Completing the 20-Element Periodic Table is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **3.7: Completing the 20-Element Periodic Table** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.3.8: The Brief Periodic Table is Complete

With element number 20, all of the elements required for the abbreviated periodic table have been described. As noted above, the placement of electrons in elements with atomic number 21 and higher is a little too complicated to explain here. However, these elements are important and they are all shown in the complete periodic table at the end of this chapter. Among the heavier elements in the complete periodic table are the transition metals, including the important metals chromium, manganese, iron, cobalt, nickel, and copper. Also included are the lanthanides and the actinides. Among these elements are thorium, uranium, and plutonium, which are important in nuclear energy and nuclear weaponry.

The abbreviated periodic table with the first 20 elements is illustrated in Figure 3.9. In addition to atomic number and atomic mass, this table shows the Lewis symbol of each element. It is seen that the symbols of the elements in the same vertical columns have the same number of dots showing identical configurations for their outer-shell electrons. This very simple, brief table contains much useful information, and it is recommended that the reader become familiar with it and be able to reproduce the Lewis symbols for each of the 20 elements. As will be seen in later chapters, the chemistry of the first 20 elements tends to be straightforward and easily related to the atomic structures of these elements.

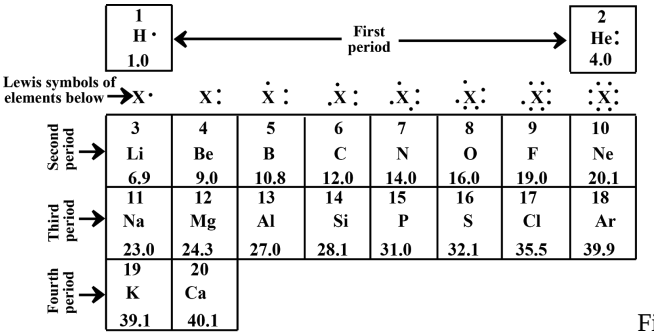


Figure 3.9. Abbreviated 20-element version of the periodic table showing Lewis symbols of the elements.

In examining the periodic table, hydrogen should be regarded as having unique properties and not belonging to a specific group. Otherwise, the elements in vertical columns belong to *groups* with similar chemical properties. Excluding hydrogen, the elements in the first group on the left of the table — lithium, sodium, and potassium — are *alkali metals*. In the elemental state alkali metals have a very low density and are so soft that they can be cut with a knife. Freshly cut, an alkali metal surface has a silvery-white color which almost instantaneously turns to a coating of gray metal oxide with exposure to air. The alkali metals (represented by M, below) react violently with water,

$$2M + 2H2O \rightarrow 2MOH + H2 \quad (3.5.4)$$

to produce the metal hydroxides, strongly basic substances that can be very destructive to flesh that they contact. The alkali metals react with elemental chlorine to produce the ionic chloride salts including, in addition to NaCl shown in Figure 3.8, LiCl and KCl. The second group of the abbreviated periodic table contains beryllium, magnesium, and calcium, all known as *alkaline earth* metals. Freshly exposed surfaces of these metals have a grayish-white luster. These metals are highly reactive to form doubly charged cations (Be2+, Mg2+, Ca2+) by the loss of 2 electrons per atom. The second group from the right, which in the abbreviated periodic table consists of fluorine and chlorine, is known as the **halogens**. These elemental halogens are diatomic gases in which the two atoms of F2 or Cl2 are held together by a single covalent bond consisting of two shared electrons. These elements are the most nonmetallic of the elements. Rather than losing electrons to produce positively charged cations, as is common with metals, the halogens readily gain electrons to complete their outer shell electron octets, producing F- and Cl- anions. The far right group of the abbreviated periodic table is composed of the *noble gases*, helium, neon, and argon. These elements have complete outer shells, exhibit no tendency to enter into chemical bonds, and consist of individual gas-phase atoms.

This page titled 2.5.3.8: The Brief Periodic Table is Complete is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **3.8: The Brief Periodic Table is Complete** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.3.E: The Elements - Basic Building Blocks of Green Chemicals (Exercises)

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. Match the law or observation denoted by letters below with the portion of Dalton's atomic theory that explains it denoted by numbers:

A. Law of Conservation of Mass B. Law of Constant Composition

C. The reaction of C with O_2 does not produce SO_2

D. Law of Multiple Proportions

1. Chemical compounds are formed by the combination of atoms of different elements in definite constant ratios that usually can be expressed as integers or simple fractions

2. During the course of ordinary chemical reactions, atoms are not created or destroyed

3. During the course of ordinary chemical reactions, atoms are not changed to atoms of other elements

4. Illustrated by groups of compounds such as $CHCl_3$, CH_2Cl_2, or CH_3Cl

2. Explain why it is incorrect to say that atomic mass is the mass of any atom of an element. How is atomic mass defined?

3. Define what is meant by the notation ${}^{y}_{x}A$? What do y, x, and A mean?

4. What is the Lewis symbol of hydrogen and what does it show? What is the Lewis formula of H2 and what does it show?

5. Why should hydrogen be considered in a separate category of the periodic table?

6. Consider the Lewis symbol of helium and explain how the helium atom illustrates the concepts of electron shell, filled electron shell, and noble gases.

7. What does helium have to do with cyrogenics?

8. Use three dots to show all the electrons in the lithium atom, Li. What does this show about inner and outer electrons and why Li produces Li+ cation?

9. In what respect may it be argued that beryllium is definitely not a green element?

10. What are two elemental oxygen species other than molecular O_2 found at very high altitudes in the stratosphere? How do they get there?

11. In what respects may carbon be classified as the "element of life"?

12. How are a specific kind of fluorine compounds related to stratospheric ozone? What does this have to do with green chemistry?

13. How does neon illustrate important points about the octet and the octet rule?

14. Of the following, the **untrue** statement pertaining to matter, atoms, and elements is

A. All matter is composed of only about a hundred fundamental kinds of matter called elements.

B. Each element is made up of very small entities called atoms

C. All atoms of the same element have the same number of protons and neutrons and the same mass

D. All atoms of the same element behave identically chemically

E. All atoms of the same element have the same number of protons

15. Given that in the periodic table elements with atomic numbers 2, 10, and 18 are gases that do not undergo chemical reactions and consist of individual molecules, of the following, the statement most likely to be true is

A. Elements with atomic numbers 3, 11, and 19 are also likely to be gases.

B. Elements with atomic numbers 3, 11, and 19 would not undergo chemical reactions.

C. Elements with atomic numbers 10 and 18 would be at opposite ends of the table.

D. The element with atomic number of 11 may well be a highly reactive metal.

E. The properties of elements with atomic number 3, 11, and 19 would have chemical properties that are much different from each other.

16. The two atoms represented below

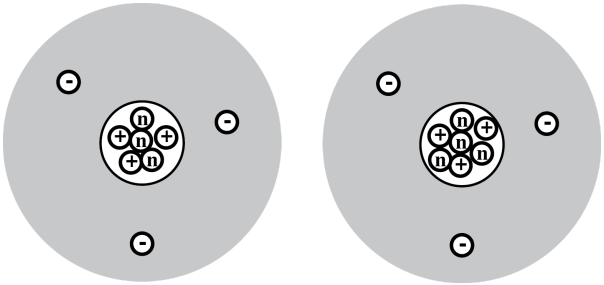
A. Are of different elements.

B. Are atoms of deuterium, a form of hydrogen, H.

C. Are of the same element.

D. Are **not** isotopes of the same element.

E. Are of two elements with atomic numbers 6 and 7.



17. Of the following, the statement that is **untrue** regarding chemical bonding and compounds is

A. Chemical bonds occur only in compounds, not in pure elements.

B. Molecules of H2 are held together by bonds consisting of shared electrons.

C. Ionic compounds consist of charged atoms or groups of atoms.

D. Both pure elemental hydrogen, H_2, and the compound water, H_2O, have covalent bonds.

E. An atom that has more electrons than protons is an anion.

18. Suggest a material that is a source of electrons in a fuel cell used to generate electricity. What may accept electrons?

19. What are semiconductors? What is the most important semiconductor discussed in this chapter.

20. What is the most notable chemical characteristic of elemental oxygen?

21. What are some reasons that aluminum can be regarded as a green metal?

22. What are some of the toxic elements or elemental forms among the first 20 elements?

23. What is a common air pollutant gas that contains sulfur?

24. Why does the abbreviated periodic table stop at atomic number 20?

25. Suggest why calcium might be particularly important in the diet of (A) children and (B) older people?

26. Which elements among the first 20 are commonly present in fertilizers used to enhance the growth of food crops?

27. What is the special significance of the carbon isotope with 6 neutrons in its nucleus?

28. What is the single exception to the rule that all atoms contain at least 1 neutron?

29. What is the single exception to the rule that noble gases contain stable octets of electrons?

30. What is the outer-shell electron configuration of metals? What does this have to do with their chemical behavior?

31. What is it about the carbon atom that enables millions of organic compounds to exist?

32. What are some of the forms of elemental carbon and their uses? Which of these was discovered only relatively recently?

33. What is the major chemical characteristic of elemental nitrogen? What is a major advantage afforded by this characteristic? In what respect is this a problem?

34. What are two applications that elemental magnesium and aluminum have in common?

35. How do copper and silica differ in the way that they transfer communications signals?

36. Using the octet rule, propose a Lewis formula for O_2.

This page titled 2.5.3.E: The Elements - Basic Building Blocks of Green Chemicals (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **3.E: The Elements - Basic Building Blocks of Green Chemicals (Exercises)** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.3.R: References

1. Manahan, Stanley E., *Fundamentals of Sustainable Chemical Science*, Taylor & Francis/CRC Press, Boca Raton, FL, 2009.

2. Brown, Theodore L., H. Eugene Lemay, Bruce E. Bursten, Catherine J. Murphy, and Patrick Woodward, *Chemistry: The Central Science*, 11th ed., Prentice Hall, Upper Saddle River, NJ, 2008.

3. Chang, Raymond, *Chemistry*, 10th ed., McGraw-Hill, New York, 2009

SUPPLEMENTARY REFERENCES

Brown, Lawrence S., and Thomas A. Holme, *Chemistry for Engineering Students*, 2nd ed., Brooks/Cole Cengage Learning, Belmont, CA, 2011.

Denniston, Katherine J., Joseph J. Topping, and Robert L. Caret, *General, Organic, and Biochemistry*, 7th ed., McGraw-Hill, New York, 2011.

Guinn, Denise, and Rebecca Brewer, *Essentials of General, Organic, and Biochemistry: An Integrated Approach*, W. H. Freeman and Co., New York, 2009.

Hill, John W., and Doris Kolb, *Chemistry for Changing Times*, 11th ed, Prentice Hall, Upper Saddle River, NJ, 2006.

McMurry, John, David S. Ballantine, Carl A. Hoeger, Virginia E. Peterson, and Mary E. Castellion, *Fundamentals of General, Organic, and Biological Chemistry*, 6th ed., Prentice-Hall, Upper Saddle River, NJ, 2009.

Mickulecky, Peter J., Katherine Brutlag, Michelle Gilman, and Brian Peterson, *Chemistry Workbook for Dummies*, Wiley, Hoboken, NJ, 2008.

Middlecamp, Catherine, Steve Keller, Karen Anderson, Anne Bentley, Michael Cann, *Chemistry in Context*, 7th ed., McGraw-Hill, Dubuque, IA, 2011.

Moore, John W., Conrad L. Stanitski, and Peter C. Jurs, *Chemistry: The Molecular Science*, 4th ed., Brooks/Cole Cengage Learning, Belmont, CA, 2011.

Smith, Janice G., *Principles of General, Organic, and Biochemistry*, McGraw-Hill, New York, 2011.

This page titled 2.5.3.R: References is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **3.R: References** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

2.5.4: Compounds- Safer Materials for a Safer World

“The millions of known chemical compounds vary tremendously in their properties. For example, some are so unstable that just touching them with a steel spatula will cause them to explode spontaneously. Others are so stable and persistent that they last practically forever in the environment”

2.5.4.1: Chemical Bonds and Compound Formation

2.5.4.2: Electrons Involved in Chemical Bonds and Octets of Electrons

2.5.4.3: Sodium Chloride and Ionic Bonds

2.5.4.4: Covalent Bonds in H2 and Other Molecules

2.5.4.5: Covalent Bonds in Compounds

2.5.4.6: Covalent Bonds and Green Chemistry

2.5.4.7: Predicting Formulas of Covalently Bound Compounds

2.5.4.8: Chemical Formulas, the Mole, and Percentage Composition

2.5.4.9: What Are Chemical Compounds Called?

2.5.4.10: Acids, Bases and Salts

2.5.4.11: Questions and Problems

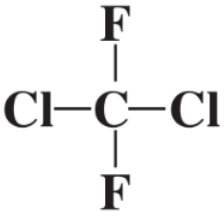
This page titled 2.5.4: Compounds- Safer Materials for a Safer World is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.4.1: Chemical Bonds and Compound Formation

Chemical compounds consist of molecules or aggregates of ions composed of two or more elements held together by chemical bonds. Several examples of chemical compounds including water (H_2O), ammonia (NH_3), and sodium chloride (NaCl) were given in earlier chapters. This chapter addresses chemical compounds in more detail, including aspects of their green chemistry.

A crucial aspect of chemical compounds consists of the kinds of bonds that hold them together. As noted earlier, these may be covalent bonds composed of shared electrons or ionic bonds consisting of positively charged cations and negatively charged anions. The strengths of these bonds vary and are important in determining compound behavior. For example, chlorofluorocarbons, such as dichlorodifluoromethane, Cl_2CF_2, are so stable that they persist in the atmosphere and do not break down until reaching very high altitudes in the stratosphere, where the release of chlorine atoms destroys stratospheric ozone. The extreme stabilities of the chlorofluorocarbons are due to the very high strengths of the C-Cl and C-F bonds by which chlorine and fluorine are bonded to a central carbon atom. The proper practice of green chemistry requires that substances that get released to the environment break down readily. Since Cl_2CF_2 is so stable when released to the atmosphere, it cannot be regarded as being a very good green chemical.

Another important aspect of the way in which chemical compounds are put together is molecular structure, which refers to the shape of molecules. Consider the Cl_2CF_2 compound just mentioned in which the Cl and F atoms are bonded to a single carbon atom. To represent this molecule as the flat structure (below)



is not totally correct because not all of the 5 atoms in the compound lie in the same plane. Instead, the F and Cl atoms can be visualized as being distributed as far apart as possible in three dimensions around a sphere, at the center of which is the C atom. This can be represented as



where the two Cl atoms are visualized as being above the plane of the book page toward the reader and the two F atoms are represented as being below the plane of the page away from the reader. The shape of molecules is very important in determining the ways in which they interact with other molecules. For example, the molecules of enzymes that enable metabolism to occur in living organisms recognize the substrate molecules upon which they act by their complementary shapes.

What Are Green Chemical Compounds?

Chemical compounds vary markedly in the degree to which they are "green."Dichlorodifluoromethane, Cl_2CF_2, the chlorofluorocarbon discussed above, is definitely not green. That is not because it is toxic — it is one of the least toxic synthetic compounds known — but because it is so extremely stable and persistent in the atmosphere and can cause stratospheric ozone destruction. The compounds that have replaced it, the hydrofluorocarbons and hydrochlorofluorocarbons, are much more green because they do not last long when released in the atmosphere or do not contain ozone-damaging chlorine.

There are several characteristics of compounds that meet the criteria of being green. These are the following:

• Preparation from renewable or readily available resources by environmentally friendly processes

• Low tendency to undergo sudden, violent, unpredictable reactions such as explosions that may cause damage, injure personnel, or cause release of chemicals and byproducts to the environment.

• Nonflammable or poorly flammable

• Low toxicity

• Absence of toxic or environmentally dangerous constituents, particularly heavymetals

• Facile degradability, especially biodegradability, in the environment.

• Low tendency to undergo bioaccumulation in food chains in the environment

An example of a green compound is sodium stearate, common hand soap. This common substance is prepared by reacting byproduct animal fat with sodium hydroxide, which is prepared by passing an electrical current through saltwater. Flushed down the drain, sodium stearate reacts with calcium in water to form a solid, calcium stearate, the white solid that composes “bathtub ring.” This removes the soap from the water and the nontoxic calcium stearate readily undergoes biodegradation so that it does not persist in the environment.

This page titled 2.5.4.1: Chemical Bonds and Compound Formation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **4.1: Chemical Bonds and Compound Formation** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.4.2: Electrons Involved in Chemical Bonds and Octets of Electrons

The electrons in the *outermost shell* of atoms are those that become involved in chemical bonds. These are called **valence electrons**. Refer back to the *Lewissymbols* of the elements shown in Figure 3.9. Note that the three elements on the right of the table are *noblegases* that are chemically content with their filled outer electron shells containing 2 electrons in the case of helium and 8 each for neon and argon. As a basis for the understanding of chemical bonds consider that the other elements tend to attain the filled electron shells of their nearest-neighbor noble gases by sharing, losing, or gaining electrons. Of these elements, the only one that we will consider in detail that attains a helium-like electron configuration is hydrogen, H, each atom of which almost always has access to 2 electrons shared in covalent bonds. The other elements that we will consider, carbon and higher, attain 8 electrons in their outer shells by chemical bonding. This is the basis of the **octet rule**, the tendency of atoms to attain stable outer shells of 8 electrons by forming chemical bonds. The octet rule is immensely useful in explaining and predicting chemical bonds and the formulas and structures of chemical compounds and will be emphasized in this chapter.

Some examples of the kinds of bonding arrangements discussed above have already been illustrated in Chapter 3. Figure 3.1 illustrates that, even in the elemental form, H2, hydrogen atoms have 2 valence electrons in the diatomic molecule. Examples of elements that have 8 valence electrons as the result of chemical bonding were also shown. Figure 3.6 illustrates the two N atoms in the N_2 molecule sharing 6 electrons in a covalent bond so that each of the atoms may have an octet. Figure 3.8 shows that 2 Cl atoms, each with 7 valence electrons, share 2 electrons in the covalent bond of the Cl_2 molecule to attain octets. The same figure shows that Na loses its single valence electron in forming ionic NaCl to produce the Na^+ ion, which has an octet of electrons in its outer shell. In forming the same ionic compound, Cl gains an electron to become the Cl^- anion, which also has a stable octet of outer-shell electrons.

In the remainder of this chapter, the octet rule will be used in explaining the formation of chemical compounds consisting of two or more different elements bonded together. It was already used to show the bonding in ionic sodium chloride in Figure 3.8. One of the best compounds for showing the octet rule in covalent compounds is methane, CH_4, shown in Figure 4.1. The molecule of CH_4 is produced when an atom of carbon with 4 outer electrons (see Figure 3.9) attains an octet of 8 electrons by sharing with H atoms.

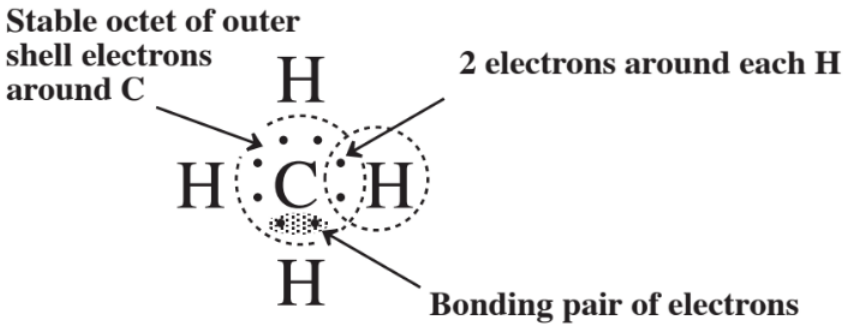


Figure 4.1. Stable outer electron shells from covalent bonding in compounds

Each H atom has 1 electron to donate to the sharing arrangement, so by each of 4 H atoms contributing an electron the carbon atom can gain an octet. Each of the H atoms has access to 2 electrons in the single covalent bond that connects it to the C atom. Examination of Figure 4.1 implies that the 4 H atoms and the C atom all lie in the same plane in a flat structure. But that is not the case because of the tendency for the 4 electron pairs composing the 4 covalent bonds to be oriented as far as possible from each other around the sphere of the carbon atom. The resulting structure is a little hard to illustrate on paper, but one way to approximate it is with a ball and stick model that represents the atoms as balls and the chemical bonds as sticks connecting the atoms. Figure 4.2 is an illustration of the ball and stick model of CH_4.

This page titled 2.5.4.2: Electrons Involved in Chemical Bonds and Octets of Electrons is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **4.2: Electrons Involved in Chemical Bonds and Octets of Electrons** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.4.3: Sodium Chloride and Ionic Bonds

Many atoms and groups of atoms in chemical compounds are **ions** that have an electrical charge because of their unequal numbers of protons and electrons. **Cations** are positively charged ions and **anions** are negatively charged ions. Compounds consisting of ions are **ioniccompounds** and the bonds holding them together are **ionic bonds**. Ionic bonds depend upon the mutual attraction between positive cations and negative anions for their bond strength (oppositely charged bodies attract each other, whereas negatively charged bodies repel each other).

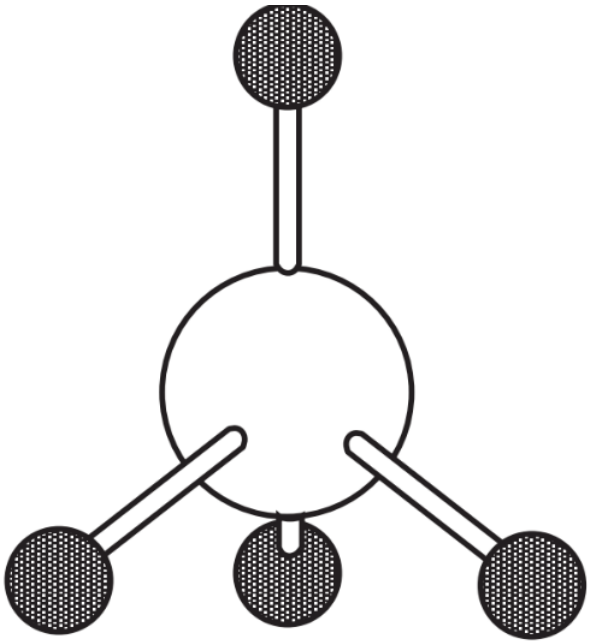


Figure 4.2. Ball and stick model of CH_4 in which 4 H atoms (shaded balls) are arranged in three dimensions around a carbon atom.

The formation of ions based upon the octet rule is readily seen for the well-known ionic compound, sodium chloride, NaCl, as illustrated in Figure 4.3. By losing an electron to become the Na^+ cation, sodium's underlying shell of 8 electrons becomes the ion's outer shell with a stable octet. Chlorine attains a stable octet of 8 outer-shell electrons by gaining 1 electron per atom to produce Cl^- ion.

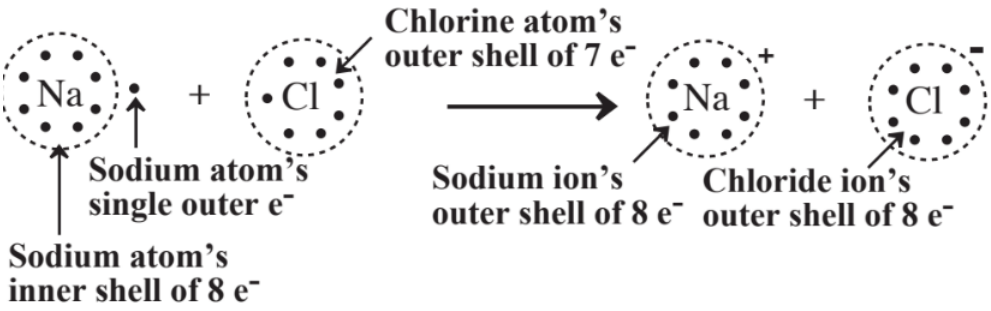


Figure 4.3. Formation of ionic sodium chloride from neutral Na and Cl atoms.

Sodium chloride is a very stable compound because of the mutual attraction of oppositely charged ions. But the ions have to be arranged in an optimum manner for this attraction to be effective. Since oppositely charged ions attract each other, but ions with the same charge are mutually repulsive, the ions in an ionic compound such as sodium chloride have to be packed to maximize attraction and minimize repulsion. The arrangement that does this for NaCl is shown by a ball and stick model in Figure 4.4.

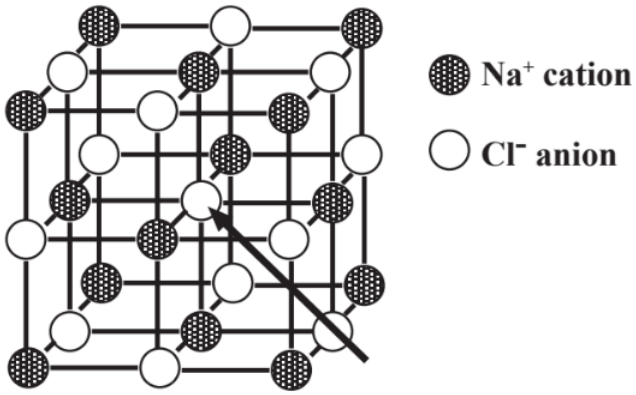


Figure 4.4. Representation of the structure of NaCl. This structure can be visualized in three dimensions as “balls” with interconnecting “sticks.” The balls are Na^+ and Cl^- ions. The closest neighbors to the Cl^- ion marked with an arrow are 6 Na^+ ions. The 6 closest neighbors to the Cl^- ion marked with an arrow are 6 Na^+ ions.

Although it may be a little hard to imagine for a model represented on paper, the six nearest neighbors of each negatively charged Cl^- anion are Na^+ cations. And the six nearest neighbors of each positively charged Na^+ cation are negatively charged Cl^- anions

In reality, ions are more accurately represented in an ionic structure as spheres that touch. The Na^+ cation is significantly smaller than the Cl^- anion, so a more accurate representation of NaCl than that shown in Figure 4.4 would show rather large Cl- spheres between which are nestled barely visible Na+spheres. But the imperfect ball and stick model shown in Figure 4.4 shows several important points regarding ionic NaCl. It illustrates the relative positions of the ions. These positions, combined with ionic charge and size, determine the **crystalstructure** of the solid crystal of which the ionic compound is composed. Furthermore, examination of the figure shows that no single Cl^- anion belongs to a specific Na^+ cation, and no single Na^+ cation belongs to a specific Cl^- anion. So, although the chemical formula NaCl accurately represents the relative numbers of Na and Cl atoms in sodium chloride it *doesnot* imply that there are discrete molecules consisting of 1 Na and 1 Cl. For this reason it is not correct to refer to a *molecule* of sodium chloride because distinct molecules of ionic compounds do not exist as such. Instead, reference is made to **formulaunits** of ionic compounds, where a formula unit of NaCl consists of 1 Na^+ ion and 1 Cl^- ion, the smallest quantity of a substance that can exist and still be sodium chloride.

The stabilities of chemical compounds are all about energy. In general, the more energy released when a compound forms, the more stable the compound is. Sodium chloride could be formed by reacting elemental solid sodium with elemental Cl_2 gas,

$$2\,\mathrm{Na(solid)} + \mathrm{Cl_2\,(gas)} \rightarrow 2\,\mathrm{NaCl\ (solid)}$$

to produce solid sodium chloride. This reaction releases a large amount of energy and elemental sodium burns explosively in chlorine gas. The reaction can be viewed in terms of the following steps.

1. The atoms in solid Na are taken apart, which requires energy.
2. Each molecule of Cl_2 is taken apart, which requires energy.
3. An electron is taken from each Na atom to produce Na^+ ion, which requires energy.
4. An electron is added to each Cl atom to produce a Cl^- ion, which releases energy.
5. All the Na^+ cations and 1 Cl^- anion are assembled in a 1/1 ratio in a crystal lattice to produce NaCl, which releases a very large quantity of energy.

The very large amount of energy involved in Step 5 is called the **latticeenergy** and is primarily responsible for the high stability of ionic compounds. A general picture of the energy involved is shown in Figure 4.5.

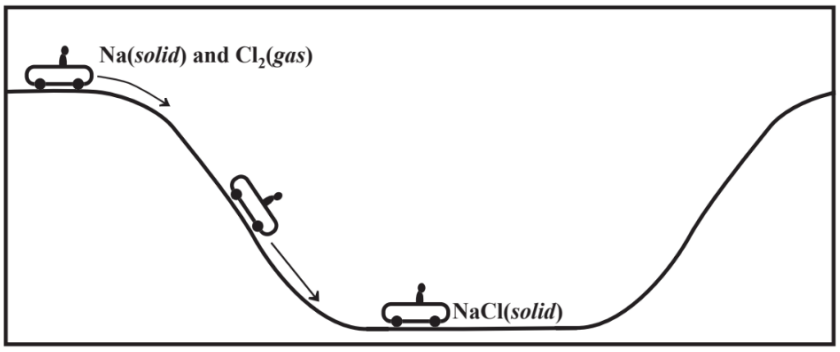


Figure 4.5. The energy involved in forming a compound is much like that involved in rolling a cart down a hill. In the case of sodium chloride, a lot of energy is released in forming solid crystalline ionic NaCl from solid Na metal and gaseous Cl_2

The differences in ionic size noted above are represented in Figure 4.6 for several monatomic (1-atom) ions from elements close to each other in the periodic table. The figure shows that negative monatomic ions are generally larger than positive monatomic ions formed from elements that are nearby in the periodic table. Thus, the negative F^- ion is larger than the positive Na+ion, although both ions have the same number of electrons (10) and the atomic number of Na is higher than that of F. It is

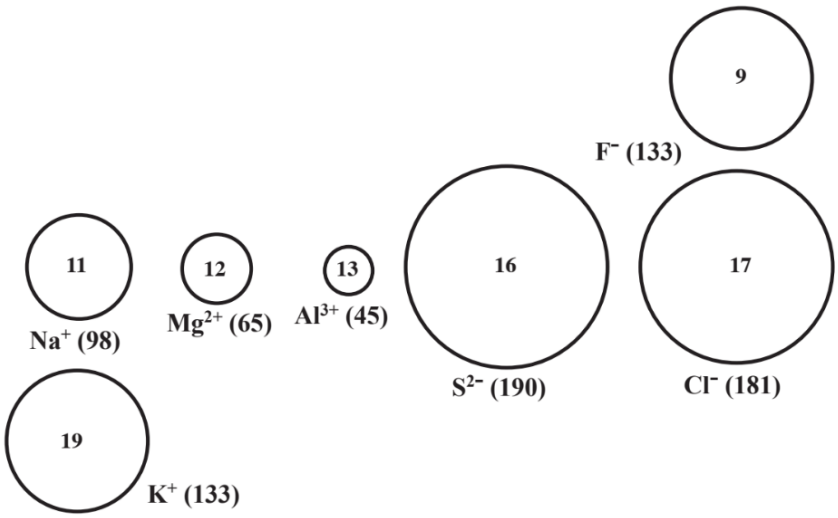
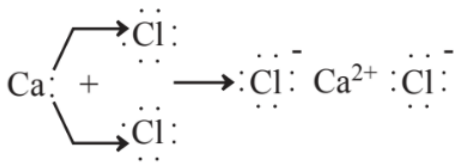


Figure 4.6. Relative sizes of several ions formed by removing electrons from or adding electrons to several elements shown in the abbreviated periodic table. The ion diameters are given in picometers (pm) in parentheses after the symbol for each ion. The atomic numbers of the elements from which each ion has been formed are shown inside the circles representing the ions.

seen that for ions in the same group of elements that have the same charge, the ion from the element with higher atomic number is larger. Figure 4.6 shows the Cl^-ion larger than the F^- ion and the K^+ ion is larger than the Na^+ ion. As electrons are removed from elements in the same period of the periodic table to produce progressively more highly charged cations, ion size shrinks notably as shown by the order of ion sizes $Na^+ > Mg^{2+} > Al^{3+}$, each of which has 10 e-. This occurs because as the charge of the nucleus becomes larger relative to the charge of the negative electron cloud around it, the cloud is drawn closer to the nucleus, shrinking ion size. As electrons are added to atoms to produce more highly charged anions, the anion size increases because more electrons occupy more space. So S^{2-} ion is larger than Cl^- ion.

In order to further understand ion formation, several more examples can be considered. Calcium and chlorine react,

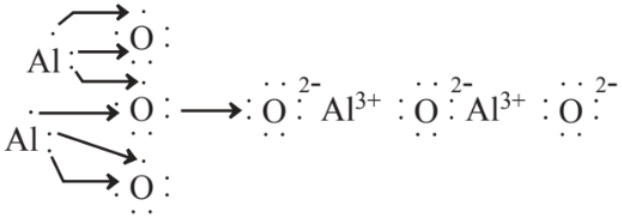


to form calcium chloride, $CaCl_2$. This compound is a byproduct of some industrial processes, from which its disposal can be a problem. It is commonly used as road salt to melt ice and snow on streets and highways. Although calcium chloride is effective in this respect, it is corrosive to automobiles and calcium chloride is a pollutant salt that can contribute to excess salt levels in bodies of water. A "greener," though more costly substitute is calcium acetate, $Ca(C_2H_3O_2)_2$. This compound is composed of Ca^{2+} ions and acetate ($C_2H_3O_2^{2-}$) anions. Its advantage is that bacteria on soil and in water readily cause biodegradation of the acetate anion as shown by the reaction,

$$Ca(C_2H_3O_2)_2 + 4\,O_2 \xrightarrow{Bacteria} CaCO_3 + 3\,CO_2 + 3\,H_2O$$

from which the calcium ends up as calcium carbonate, a common component of rock and soil.

Another example of the formation of an ionic compound is the following reaction of aluminum metal with elemental oxygen,



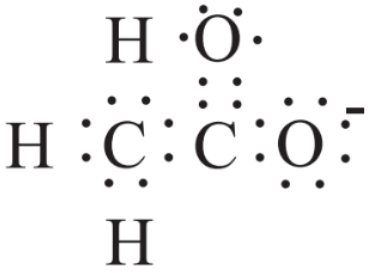
This reaction produces aluminum oxide for which the chemical formula is Al_2O_3. This compound is the source of aluminum in bauxite, the ore from which aluminum is produced and is an important industrial chemical. Called alumina, aluminum oxide itself has many applications including its use for abrasives and sandpaper, as a raw material for ceramics, and as an ingredient of antacids and antiperspirants.

Exercise: Show the ionic products of the reaction of the metal and nonmetal indicated

Metal	Nonmetal	Cation formed	Anion formed	Ionic compound
Na	Cl	(a)_____	(b)_____	(c)_____
K	O	(d)_____	(e)_____	(f)_____
Ca	Cl	(g)_____	(h)_____	(i)_____

Answers: (a) Na^+, (b) Cl^-, (c) NaCl, (d) K^+, (e) O^{2-}, (f) K_2O, (g) Ca^{2+}, (h) Cl^-, (i) $CaCl_2$

In addition to ions formed from single atoms losing or gaining electrons, many ions consist of groups of atoms covalently bound together, but having a net electrical charge because of an excessor a deficiency of electrons. An example of such an ion is the acetate ion shown above in the formula of calcium acetate, $Ca(C_2H_3O_2)_2$. The structural formula of the acetate anion, $C_2H_3O_2^-$, is shown below in which the two carbon atoms are joined with a single covalent bond consisting of two shared electrons, each of the three H atoms are joined to one of the carbon atoms by a single covalent bond and the other carbon atom is joined to one oxygen with a single covalent bond and to the other by a double covalent bond consisting of 4 shared electrons. The net charge on the ion is -1.



Ionic Liquids and Green Chemistry

Most common ionic compounds such as sodium chloride are hard solids because the ions of which they are composed are relatively small and packed tightly together in the crystalline lattice. These ionic compounds must be heated to very high temperatures before they melt, 801° C for NaCl, for example. In recent years, ionic compounds have been developed that are liquids under ordinary conditions. The ions in these **ionicliquids** are composed of large organic molecules composed of skeletons of numerous carbon atoms bonded to other atoms and having a net charge. The charges on the ions in such compounds is much less concentrated than in simple inorganic compounds like NaCl, the large ions move readily relative to each other in the ionic crystal, and the compound is liquid at low temperatures. A common example of an ionic liquid compound is decylmethylimidazolium hexafluorophosphate, which is a liquid at temperatures above 40°C, the temperature of a very hot summer's day.

There has been a lot of interest in the application of ionic liquids to green chemistry. This is because many chemical reactions including those for preparing polymers used in synthetic fabrics or pharmaceutical compounds are carried out in liquid solvents, which tend to evaporate and contaminate air and to pose disposal problems. Furthermore, although the solvent properties in chemical synthesis often play a strong role in determining the course of the synthesis, the number of available solvents is very limited. In the case of ionic liquids, however, there is a vast variety of both cations and liquids which, combined together, could give as many as a trillion (!) different ionic liquids. This versatility enables fine-tuning the properties of the ionic liquids for specialized uses in synthesis and other applications. The ionic liquids are rather easy to recycle, adding to their green character. In addition to their applications as solvents for chemical synthesis media, ionic liquids may be useful for isolating pollutants. For example, an appropriate ionic liquid may be shaken with water contaminated with toxic heavy metals, such as lead or cadmium, which bind with the ionic liquid. Since such a liquid is normally not soluble in water, it can be physically separated from the water, carrying the heavy metals with it and leaving purified water.

This page titled 2.5.4.3: Sodium Chloride and Ionic Bonds is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **4.3: Sodium Chloride and Ionic Bonds** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.4.4: Covalent Bonds in H2 and Other Molecules

Lewis symbols can be used to show how some atoms of elements on the left side of the table with only one or two outer-shell electrons can lose those electrons to form cations such as Na^+ or Ca^{2+}. It is also easily seen that atoms from groups near the right side of the periodic table can accept one or two electrons to gain stable octets and become anions such as Cl^- or O^{2-}. But, it is difficult to impossible to take more than two electrons away from an atom to form cations with charges greater than +2 or to add 3 or more electrons to form anions with charges of -3 or even more negative, although ions such as Al^{3+} and N^{3-} do exist. So atoms of elements in the middle of the periodic table and the nonmetals on the right have a tendency to share electrons in covalent bonds, rather than becoming ions.

It is readily visualized how mutually attracting ions of opposite charge are held together in a crystalline lattice. Shared electrons in covalent bonds act to reduce the forces of repulsion between the positively charged nuclei of the atoms that they join together. That is most easily seen for the case of the hydrogen molecule, H_2. The nuclei of H atoms consist of single protons, and the two H atom nuclei in the H_2 molecule repel each other. However, their 2 shared electrons compose a cloud of negative charge between the two H nuclei, shielding the nuclei from each other's repelling positive charge and enabling the molecule to exist as a covalently bound molecule (Figure 4.7).

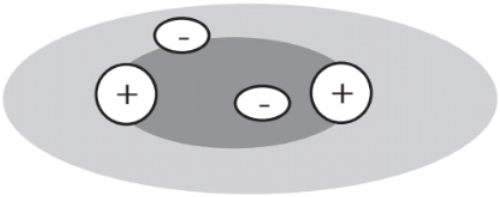


Figure 4.7. Electrons in motion between and around mutually repelling H nuclei shield the nuclei from each other constituting a single covalent bond that holds the H_2 molecule together

This page titled 2.5.4.4: Covalent Bonds in H2 and Other Molecules is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

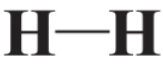
- **4.4: Covalent Bonds in H2 and Other Molecules** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.4.5: Covalent Bonds in Compounds

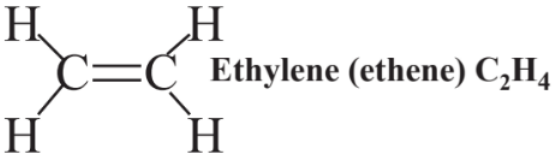
Consider next some example covalent bonds between atoms of some of the lighter elements. These are best understood in reference to Figure 3.9, the abbreviated version of the periodic table showing the Lewis symbols (outer shell valence electrons) of the first 20 elements. As is the case with ions, atoms that are covalently bonded in molecules often have an arrangement of outer shell electrons like that of the noble gas with an atomic number closest to the element in question. It was just seen that covalently bonded H atoms in molecules of H_2 have 2 outer shell electrons like the nearby noble gas helium. For atoms of many other elements, the tendency is to acquire 8 outer shell electrons— an octet— in sharing electrons through covalent bonds. This tendency forms the basis of the *octetrule* discussed in Section 4.2. In illustrating the application of the octet rule to covalent bonding Section 4.7 considers first the bonding of atoms of hydrogen to atoms of elements with atomic numbers 6 through 9 in the second period of the periodic table.

These elements are close to the noble gas neon and tend to attain a “neon-like” octet of outer shell electrons when they form covalently bonded molecules.

Covalent bonds are characterized according to several criteria. The first of these is the number of electrons involved. The most common type of covalent bond consists of 2 shared electrons and is a *single bond*. Four shared electrons as shown for the bond joining an O atom to one of the C atoms in the structure of the acetate anion above constitute a *double covalent bond*. And 6 shared electrons as shown for the very stable covalent bond joining the two N atoms in the N_2 molecule illustrated in Chapter 3, Figure 3.6 make up a *triple covalent bond*. These bonds are conventionally shown as lines in the structural formulas of molecules (large numbers of dots in a formula can get a little confusing). So the single covalent bond in H_2 is shown as



The double bond consisting of 4 shared electrons holding the two carbon atoms together in C_2H_4 (ethylene, a hydrocarbon used to make polyethylene plastic) are shown by the following:



And the very strong triple bond joining the two N atoms in the N_2 molecule are shown by three lines:

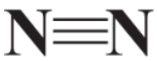


Figure 3.6 that each N atom has a pair of electrons that are not part of any chemical bonds. These are omitted from the structure above, and are not ordinarily shown when bonds are represented by lines.

Covalent bonds have a characteristic **bond length**. Bond lengths are of the general magnitude of the size of atoms, so they are measured in units of picometers (pm). The H-H bond in H_2 is 75pm long.

A third important characteristic of bonds is **bond energy**. Bond energy is normally expressed in kilojoules (kJ) required to break a mole (6.02×10^{23}) of bonds. (See Section 4.8 for a detailed definition of the mole.) The bond energy of the H-H bond in H_2 is equal to 435 kJ/mole. This means that an amount of energy required to break all the H-H bonds in a mole of H_2 (2.0 grams of H_2, 6.02×10^{23} molecules) is a very substantial 435 kJ.

This page titled 2.5.4.5: Covalent Bonds in Compounds is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **4.5: Covalent Bonds in Compounds** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.4.6: Covalent Bonds and Green Chemistry

The nature of covalent bonds is strongly related to green chemistry. Some reasons why this is so include the following:

- High-energy bonds in raw materials require a lot of energy and severe conditions, such as those of temperature and pressure, to take apart in synthesizing
- Especially stable bonds may make substances unduly persistent in the environment.
- Relatively weak bonds may allow molecules to come apart too readily, contributing to reactive species in the atmosphere or in biological systems.
- Unstable bonds or arrangements of bonds may lead to excessive reactivity in chemicals making them prone to explosions and other hazards.
- Some arrangements of bonds contribute to chemical toxicity.

An example of a substance that has a very high bond stability making it an energy-intensive source of raw material is N_2. As mentioned earlier, large amounts of energy and severe conditions are required to take this molecule apart in the synthesis of ammonia, NH_3, the compound that is the basis for most synthetic nitrogen compounds. As discussed with nitrogen in Section 2.5, *Rhizobium* bacteria growing on the roots of leguminous plants such as soybeans convert N_2 to chemically fixed nitrogen. The amount of nitrogen fixed by this totally green route is certainly a welcome addition to the pool of fertilizer nitrogen.

An example of a compound in which especially stable bonds contribute to persistence and ultimate environmental harm is provided by dichlorodifluoromethane, Cl_2CF_2, a chlorofluorocarbon implicated in the destruction of stratospheric ozone (see Chapter 3, Section 3.5, and Chapter 10). The chemical bonds in this compound are so strong that nothing attacks them until the molecule has penetrated many kilometers high into the stratosphere where extremely energetic ultraviolet radiation breaks the C-Cl bond in the molecule. This produces Clatoms that bring about the destruction of protective stratospheric ozone.

A somewhat opposite condition occurs in the case of atmospheric nitrogen dioxide, NO_2, near ground level in the atmosphere. Here the NO bond is relatively weak so that the relatively low-energy ultraviolet radiation ($h\nu$) that is close to the wavelength of visible light and penetrates to the atmosphere at or near ground level can break apart NO_2, molecules:

$$NO_2 + h\nu \rightarrow NO + O$$

The O atoms released are very reactive and interact with pollutant hydrocarbons, such as those from gasoline, in the atmosphere resulting in the disagreeable condition of photochemical smog.

Some bonding arrangements are notable for instability. These include bonds in which two N atoms are adjacent or very close in a molecule and are bonded with double bonds. Also include dare arrangements in which N and O atoms are adjacent and double bonds are also present.

The presence of some kinds of bonds in molecules can contribute to their biochemical reactivity and, therefore, to their toxicity. An organic compound with one or more C=C double bonds in the molecule is often more toxic than a similar molecule without such bonds.

By avoiding generation, use, or release to the environment of compounds with the kinds of bonds described above as part of a program of green chemistry, the practice of chemistry and the chemical industry can be made much safer. Green chemistry also considers bonds that may have to be protected in chemical synthesis. Often steps must be added to attach protecting groups to bonding groups to prevent their reacting during a step of a synthesis. After the desired step is performed, the protecting groups must be removed to give the final product. Materials are consumed and byproducts are generated by these steps, so the practice of green chemistry attempts to avoid them whenever possible.

This page titled 2.5.4.6: Covalent Bonds and Green Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **4.6: Covalent Bonds and Green Chemistry** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.4.7: Predicting Formulas of Covalently Bound Compounds

It is often possible to predict the formulas of molecules using the Lewis symbols of the elements in the compound with the octet rule for chemical bonding. This is shown in Figure 4.8 for the hydrogen compounds of several elements in the second period of the periodic table.

The prediction of chemical bonds in compounds in which H is bonded to another atom is very simple because each H atom has to be involved in sharing two electrons and the other kind of atom normally has to have a total of 8 electrons in its valence shell octet; these may be shared in bonds or present as unshared pairs. As an example, consider a well-known compound of carbon, carbon dioxide, chemical formula CO_2. The Lewis symbol of C and those of the two O atoms can be used to deduce the Lewis formula of CO_2 as shown in Figure 4.9.

As another example of the application of the octet rule, consider hydrogen peroxide, H_2O_2. This compound's formula looks a lot like that of water, but it is a lot different from water. Hydrogen peroxide decomposes to release oxygen:

$$2\,H_2O_2\ (\text{liquid}) \rightarrow O_2\,(\text{gas}) + 2\,H_2O(\text{liquid})$$

As a liquid in the form of a concentrated aqueous solution, hydrogen peroxide provides a source of oxygen so potent that it has been used in rockets. It was the treacherous oxidant used along with hydrazine (N_2H_4) fuel in the German Luftwaffe's Messerschmidt 163 rocket plane at the end of World War II. Trailing an exhaust of lethal NO_2 gas, this minuscule manned missile (on the rare occasions when it worked according to plan) was propelled rapidly into the lower stratosphere, then glided down through waves of Allied bombers, attempting to nick them with machine gunfire as it plummeted back to Earth. Few bombers were damaged but many Me-163 pilots died in the attempt, some as the result of explosions, fires, and spills from the hydrogen peroxide oxidant. Hydrogen peroxide decomposed over a catalyst was also used as a source of oxygen for the diesel engines on several German submarines near the end of World War II. Pre-dating nuclear submarines, these potentially deadly craft were the first true submersibles.

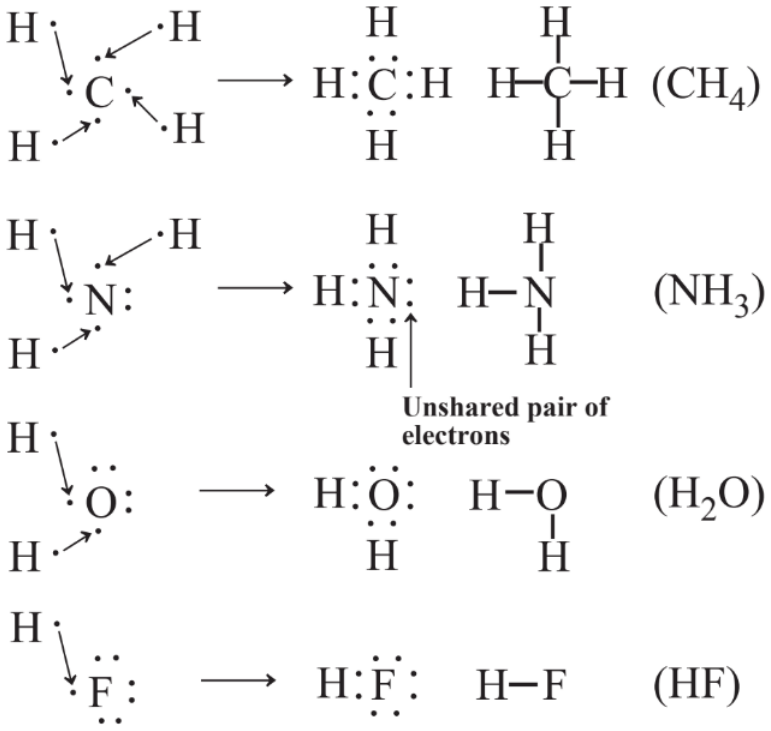
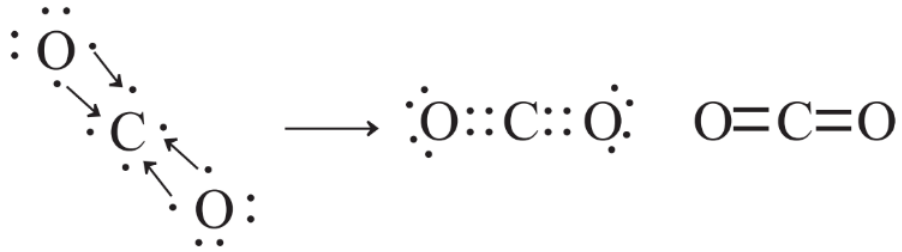


Figure 4.8. Formation of covalent hydrogen compounds with elements in the second row of the periodic table. All of carbon's valence electrons are involved in bonds in CH_4, NH_3 has 1 unshared pair of valence electrons, H_2O has 2 unshared pairs of valence electrons, and HF has 3 unshared pairs of valence electrons. The unshared pairs of electrons are not shown in the structures in which the bonds are represented by straight



lines. Figure 4.9. Bonding of two oxygen atoms to a central carbon atom to produce carbon dioxide, CO_2. In order to provide an octet of electrons around each of the three atoms with the total of only 16 valence electrons available, it is necessary to have double bonds between C and each of the O atoms.

In assembling the structure of the hydrogen peroxide molecule, one has simply to work with two O atoms each contributing 6 valence electrons and two H atoms each with 1 valence electron. The Lewis formula of the H_2O_2 molecule is



showing that all of the 14 total valence electrons are involved in chemical bonds and both oxygens have octets of outer-shell electrons.

Despite the evil nature of concentrated solutions of hydrogen peroxide, it can be regarded as a green compound in more dilute solutions, such as the 3% hydrogen peroxide commonly used to kill bacteria in treating wounds. Among its green applications, dilute hydrogen peroxide makes an effective and safe bleaching agent that is much safer to handle than elemental chlorine commonly used for bleaching and that does not produce the potentially toxic byproducts that chlorine generates. And even though it kills bacteria, hydrogen peroxide can be pumped underground to serve as an oxidant for acclimated bacteria that attack wastes that have been placed in or seeped into underground locations

Molecules That Do Not Obey the Octet Rule

In some cases the octet rule is not obeyed. This occurs when a molecule has an uneven number of electrons so that it is impossible for each atom to have an octet (an even number) of electrons. A simple example of this is nitric oxide, NO, made from an atom of N with 5 valence electrons and one of O with 6 valence electrons. The resulting molecule is shown in Figure 4.10. Since the uneven number of 11 valence electrons cannot provide complete octets of electrons around both the N and O atoms simultaneously, the NO molecule is shown in two forms in which one atom has 8 valence electrons and the other has 7. These are known as **resonance structures.**



Figure 4.10. Formation of NO. The uneven number of electrons cannot accommodate complete octets around both atoms simultaneously, so the structure is a hybrid resonance structure of the two forms shown.

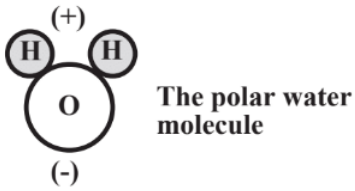
Unequal Sharing of Electrons

The Lewis formula for water,



indicates that the molecule is not symmetrical and the two H atoms are located on one side of the molecule and the O atom on the other side. One might think that the electrons shared between Hand O are shared equally. But such is not the case because the

relatively larger O atom nucleus with its 8 protons has a stronger attraction for the electrons than do the two H atom nuclei, each with only 1 proton. So the shared electrons spend relatively more time around the O atom and less around the H atoms. This gives each H atom a partial positive charge and the O atom a partial negative charge. An unequal distribution of charge such as that makes a body **polar** and the O-H bonds are **polar covalent bonds**. Because of this phenomenon, the whole water molecule is polar and can be represented as the following, where the small spheres stand for H atoms and the large one for the O atom:

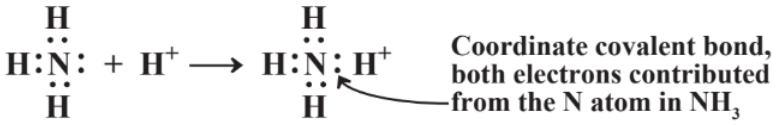


The polar water molecule

The polar nature of the water molecule has a lot to do with water as a solvent and how it behaves in the environment and in living systems. These aspects are discussed in more detail in Chapter 8.

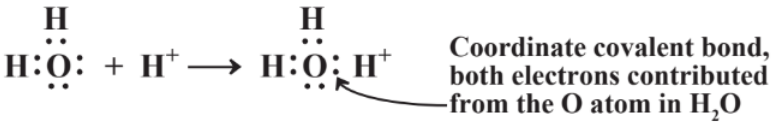
When Only One Atom Contributes to a Covalent Bond

In some cases only one of the two atoms joined by a covalent bond contributes both the electrons in the bond. This occurs with ammonia, NH_3, dissolved in water. Water contains dissolved H^+ cations, and the more acidic water is, the higher the concentration of H^+. The H^+ cation, would be stabilized by two electrons which it can get by binding with dissolved NH_3 as shown in Reaction 4.7.2. Both of the electrons shared between N and the H^+ cation now bound to it as part of a new species, the ammonium ion, NH_4^+, were contributed by the N atom. Such a covalent bond is called a **coordinate covalent bond** or a **dative bond**. In the case of NH_4^+, once the coordinate covalent N-H bond is formed, it is indistinguishable from all the other N-H bonds.



The formation of the coordinate covalent bond in NH_4^+ is very useful when soil is fertilized with nitrogen. The most economical way to apply nitrogen fertilizer is by injecting NH_3 into the soil, but NH_3 is a gas that would be expected to rapidly evaporate from soil. Instead, it becomes attached to H^+ ion from the water in the soil and is bound to the soil as the NH_4^+ ion.

Another important example of a coordinate covalent bond occurs in water. As discussed in Section 4.9, acids, which are very important materials commonly dissolved in water, produce the hydrogen ion, H^+, in water. This ion does not exist simply dispersed in water. Instead, it binds strongly to a water molecule to produce the hydronium ion, H_3O^+:



This page titled 2.5.4.7: Predicting Formulas of Covalently Bound Compounds is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **4.7: Predicting Formulas of Covalently Bound Compounds** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.4.8: Chemical Formulas, the Mole, and Percentage Composition

Chemical formulas represent the composition of chemical compounds. A number of chemical formulas have been shown so far including H_2O for water and NH_3 for ammonia. A chemical formula of a compound contains a lot of significant information as shown in Figure 4.11. Included is the following:

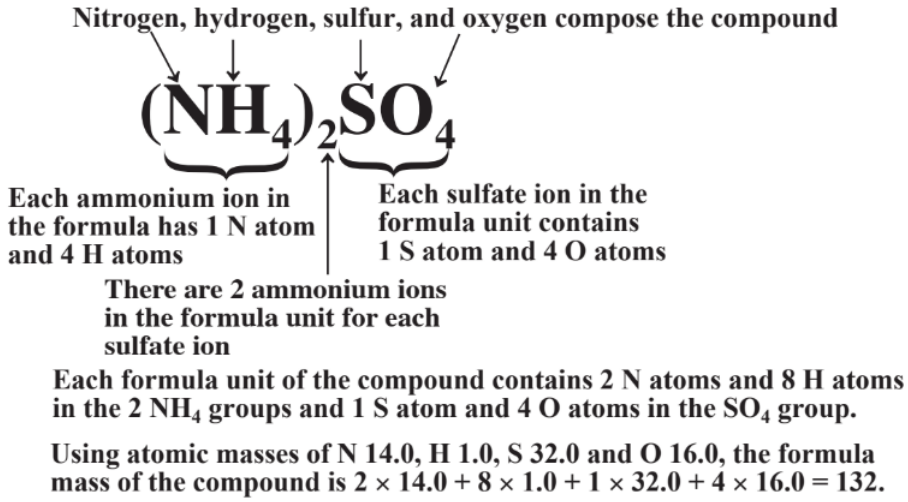


Figure 4.11. Information contained in a chemical formula

- The elements that compose the compound
- The relative numbers of each kind of atom in the compound
- How the atoms are grouped, such as in ions (for example, SO_4^{2-}) present in the compound
- With a knowledge of atomic masses, the molar mass of the compound
- With a knowledge of atomic masses, the percentage composition of the compound

Where the symbols of the elements represent letters in the alphabet of chemical language, the formulas of compounds represent words composed of those letters. As discussed in Chapter 5, formulas are put together in chemical equations that act as sentences in the chemical language to describe what chemical substances do.

The Mole

With a knowledge of atomic masses, the percentage composition of a compound is readily calculated from its formula. Before doing such a calculation for ammonium sulfate, however, it is useful to introduce the concept of the **mole**. Chemists use the mole to express quantities of materials containing a specific number of specified entities, which may be atoms of elements, molecules of elements that exist as diatomic molecules, formula units of ionic compounds, or molecules of covalent compounds. A mole of a substance is simply the atomic mass, molecular mass, or formula mass followed by *grams*. This quantity is called the **molar mass**. The masses of a mole of several typical substances are given below:

Atoms of neon, atomic mass 20.1: 20.1 g/mol

Molecules of H_2, atomic mass 1.0, molecular mass 2.0: 2.0 g/mole

Molecules of CH_4, molecular mass 16.0: 16.0 g/mole

Formula units of ionic CaO, formula mass 56.1: 56.1 g/mol

The number of specified entities in a mole of a substance is always the same regardless of the substance. This number is very large, 6.02×10^{23}, and is called **Avogadro's number**. As examples, a mole of neon contains 6.02×10^{23} neon atoms, a mole of elemental hydrogen contains 6.02×10^{23} H_2 molecules (but $2\times6.02\times10^{23}$ H atoms), and a mole of CaO contains 6.02×1023 formula units (pairs of Ca^{2+} and O^{2-} ions) of CaO.

The calculation of the percentage composition of $(NH_4)_2SO_4$ is given below. Note that the molar mass of the compound is 132 g/mol, each mol of the substance contains 2×1 = 2 mol of N, 2×4 = 8mol of H, 1 mol of S, and 4 mol of O.

$2\ \text{mol N} \times 14.0\text{g N/mol N} = 28.0\text{g N},\ = \frac{28.0g}{132g} \times 100 = 21.2$

$8\ \text{mol H} \times 1.0\ \text{g H/mol H} = 8.0\ \text{g H},\ = \frac{8.0g}{132g} \times 100 = 6.1$

$1\ \text{mol S} \times 32.0\ \text{g S/mol S} = 32.0\ \text{g S},\ = \frac{32.0g}{132g} \times 100 = 24.2$

$4\ \text{mol O} \times 16.0\ \text{g O/mol O} = 64.0\ \text{g O},\ = \frac{64.0g}{132g} \times 100 = 48.5$

Example: Given the atomic masses Ca 40.0, C 12.0, and O 16.0, what is the percentage composition of calcium oxalate, CaC_2O_4?

Answer: 31.3% Ca, 18.8% C, 50.0% O

This page titled 2.5.4.8: Chemical Formulas, the Mole, and Percentage Composition is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **4.8: Chemical Formulas, the Mole, and Percentage Composition** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.4.9: What Are Chemical Compounds Called?

The naming of chemical compounds can get a little complicated. This is particularly true of organic compounds, the names of which are discussed in Chapter 9. Some of the simpler aspects of naming inorganic compounds are discussed here.

In naming compounds, prefixes are used to represent the relative numbers of atoms in the formula unit of the compound. These prefixes through number 10 are given below:

1-mono 3-tri 5-penta 7-hepta 9-nona

2-di 4-tetra 6-hexa 8-octa 10-deca

The first class of inorganic compounds to be addressed here are **binarymolecular compounds**. Binary molecular compounds are composed of only 2 kinds of elements and do not contain ions. For these compounds, the first part of the name is simply the name of the first element in the compound formula. The second part of the name is that of the second element in the compound formula modified to have the ending *-ide*. Prefixes are added to indicate how many of each kind of atom are present in the molecule. Consider as an example the name of N_2O_5. The name of the compound is dinitrogen pentoxide where *di* indicates 2 N atoms, *pent* indicates 5 oxygen atoms, and the second element has the *ide* ending. Other examples of this system of naming are $SiCl_4$, silicon tetrachloride; S_2F_6, disilicon hexafluoride; PCl_5, phosphorus pentachloride; and SCl_2, sulfur dichloride.

A number of compounds, including binary molecular compounds, have **common names** that have been used for so long that they are part of the chemical vocabulary. An especially common example is the name of water for H_2O; its official name is dihydrogen monoxide. Another example is dinitrogen monoxide, N_2O, usually called nitrous oxide. Aluminum oxide, Al_2O_3, is commonly called alumina and silicon oxide, SiO_2, is called silica, specific mineral forms of which are sand and quartz.

Recall that *ionic compounds* are those composed of ions that are held together by ionic bonds, rather than covalent bonds. As noted in the discussion of ionic sodium chloride in Section 4.3, ionic compounds do not consist of discrete molecules, but rather of aggregates of ions whose relative numbers make the compound electrically neutral overall. Therefore, it is not correct to refer to molecules of ionic compounds but rather to *formulaunits* equal to the smallest aggregate of ions that can compose the compound. Consider, for example, the ionic compound composed of Na^+ and SO_4^{2-} ions. Every ionic compound must be electrically neutral with the same number of positive as negative charges. For the compound in question this requires 2 Na^+ ions for each SO_4^{2-} ion. Therefore, the formula of the compound is Na_2SO_4 and a formula unit contains 2 Na^+ ions and 1 SO_4^{2-} ion. Furthermore, a mole of Na_2SO_4 composed of 6.02×10^{23} formula units of Na_2SO_4 contains $2\times6.02\times10^{23}$ Na^+ ions and 6.02×10^{23} SO_4^{2-} ions. Since the ionic charges determine the relative numbers of ions, prefixes need not be used in naming the compound and it is called simply sodium sulfate.

Exercise: Give the formulas and names of compounds formed from each cation on the top row with each anion on the bottom row, below:

1. NH_4^+ 2. Ca^{2+} 3. Al^{3+}

(A) Cl^- (B) SO_4^{2-} (C) PO_4^{3-}

Answers: 1(A) NH_4Cl, ammonium chloride; 1(B) $(NH_4)_2SO_4$, ammonium sulfate; 1(C), $(NH_4)_3PO_4$, ammonium phosphate; 2(A) $CaCl_2$, calcium chloride; 2(B) $CaSO_4$, calcium sulfate; 2(C), $Ca_3(PO_4)_2$, calcium phosphate; 3(A) $AlCl_3$, aluminum chloride; 3(B) $Al_2(SO_4)_3$, aluminum sulfate; 3(C), $AlPO_4$, aluminum phosphate.

Prefixes are used in naming ionic compounds where more than 1 cation or more than 1 anion are present in the formula unit. For example, Na_2HPO_4 in which each formula unit is composed of 2 Na^+ ions, 1 H^+ ion, and 1 PO_4^{3-} ion is called disodium monohydrogen phosphate. And KH_2PO_4 is called monopotassium dihydrogen phosphate.

This page titled 2.5.4.9: What Are Chemical Compounds Called? is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **4.9: What Are Chemical Compounds Called?** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.4.10: Acids, Bases and Salts

Other than binary molecular compounds, most inorganic compounds can be classified as acids, bases, or salts. These three categories of compounds and their names are addressed briefly here.

Acids

Acids are characterized by the H^+ ion, the presence of which in water, makes the water **acidic**. An acid either contains this ion or produces it when it dissolves in water. Sulfuric acid, H_2SO_4, is an example of a compound that contains H^+ ion. Dissolved in water, a molecule of sulfuric acid exists as 2 H^+ ions and 1 SO_4^{2-} ion. An example of a compound that is classified as acidic because it *produces* H^+ ion when it dissolves in water is carbon dioxide, which undergoes the following reaction in water solution:

$$CO_2 + H_2O \rightarrow H^+ + HCO_3^-$$

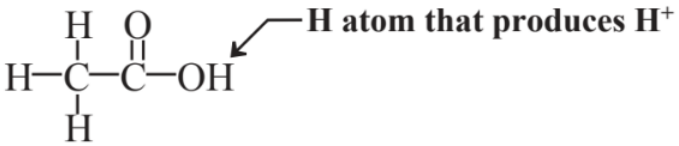
In this case, only a small fraction of the CO_2 molecules dissolved in water undergo the above reaction to produce H^+ so water solutions of CO_2 are **weakly acidic** and carbon dioxide is classified as a **weak acid**. It is the presence of dissolved CO_2 from the carbon dioxide naturally present in air that makes rainfall coming from even nonpolluted atmospheres slightly acidic and, as discussed in Chapter 9, the weakly acidic properties of CO_2 are very important in natural waters in the environment. Other acids, such as hydrochloric acid, HCl, are completely dissociated to H^+ and an anion (in the case of HCl the Cl^- anion) when they are dissolved in water; such acids are **strong acids.**

The naming of acids follows certain rules. In the case of an acid that contains only H and one other element, the acid is a hydro-ic acid. So HCl is called *hydro*chlor*ic* acid. Somewhat different rules apply when an acid contains oxygen. Some elements form acids in which the anion has different amounts of oxygen; examples are H_2SO_4 and H_2SO_3. The acid with more oxygen is an "-ic" acid, so H_2SO_4 is sulfuric acid. The acid with the lesser amount of oxygen is an "-ous" acid, so H_2SO_3 is sulfurous acid. A greater amount of oxygen than even the "-ic" acid is denoted by the prefix "per-", and a lesser amount of oxygen than even the "-ous" acid is denoted by the prefix "hypo-". These names are shown very well by the names of the oxyacids of chlorine. So the names of $HClO_4$, $HClO_3$, $HClO_2$, and HClO are, respectively, perchloric acid, chloric acid, chlorous acid and hypochlorous acid.

Acids are extremely important as industrial chemicals, in the environment, and in respect to green chemistry. About 40 million metric tons (40 billion kilograms) of sulfuric acid are produced in the United States each year. It is the number 1 synthetic chemical, largely because of its application to treat phosphate minerals to make phosphate crop fertilizers. Sulfuric acid is also used in large quantities to remove corrosion from steel, a process called steel pickling. Other major uses include detergent synthesis, petroleum refining, lead storage battery manufacture, and alcohol synthesis. About 7-8 million tons of nitric acid, HNO_3, are produced in the U.S. each year giving it a rank of 10th, and hydrochloric acid ranks about 25th with annual production around 3 million metric tons.

Acids are important in the environment. Improperly disposed acid has caused major problems around hazardous waste sites. Sulfuric acid along with smaller quantities of hydrochloric and nitric acid are the major constituents of acid rain (see Chapter 10). Acids figure prominently in the practice of green chemistry. Reclamation and recycling of acids are commonly performed in the practice of industrial ecology. As noted earlier, much of the sulfuric acid now manufactured uses a potential waste and pollutant, hydrogen sulfide, H_2S, removed from sour natural gas sources as a source of sulfur.

In cases where a relatively weak acid can be used, acetic acid made by the fermentation of carbohydrates is an excellent green alternative to stronger acids, such as sulfuric acid. Yeasts can convert the carbohydrates to ethanol (ethyl alcohol, which is present in alcoholic beverages) and other microorganisms in the presence of air convert the ethanol to acetic acid by the same process that vinegar, a dilute solution of acetic acid, is made from cider or wine. The structural formula of acetic acid is



in which only one of the 4 H atoms is ionizable to produce H+ion. The production of acetic acid is a green process that uses biological reactions acting upon renewable biomass raw materials. As a weak acid, acetic acid is relatively safe to use, and contact with humans is not usually very dangerous (we ingest dilute acetic acid as vinegar, but pure acetic acid attacks flesh and is used to remove warts from skin). Another advantage of acetic acid is that it is biodegradable, so any of it released to the environment does not persist.

Bases

A **base** either contains hydroxide ion, OH^-, or reacts with water to produce hydroxide. Most bases that contain hydroxide consist of metal cations and hydroxide; examples are sodium hydroxide, NaOH, and calcium hydroxide, $Ca(OH)_2$. The most common basic substance that produces hydroxide in water is ammonia, NH_3, which reacts with water as follows:

$$NH_3 + H_2O \rightarrow NH_4^+ + OH^-$$

Only a small fraction of the ammonia molecules undergo this reaction in water, so ammonia does not produce much OH-in water and is known as a **weak base**. The metal hydroxides, such as KOH, that completely dissociate in water are **strong bases**. Metal hydroxides are named by the metal followed by "hydroxide." Therefore, $Mg(OH)_2$ is magnesium hydroxide.

Salts

Acids and bases react to form a **salt**, an ionic compound that has a cation other than H+and an anion other than OH^-. This kind of reaction always produces water and is known as a **neutralization reaction**. The most well known salt is sodium chloride, NaCl. Although it is commonly what one means in referring to "salt," there are many other salts as well. These include calcium chloride, $CaCl_2$, used to melt road ice, sodium carbonate, Na_2CO_3, used in cleaning formulations and potassium chloride, KCl, a source of potassium fertilizer for crops. A typical neutralization reaction is the one between NaOH and hydrochloric acid, HCl, to produce sodium chloride:

$$\underbrace{NaOH}_{\text{Base}} + \underbrace{HCl}_{\text{acid}} \rightarrow \underbrace{NaCl}_{\text{a salt, sodium chloride}} + \underbrace{H_2O}_{\text{water}}$$

Salts are named very simply with just the name of the cation followed by that of the anion. The charges of the ions determine the formulas of the salts, so it is not necessary to add prefixes to denote the relative numbers of each ion. Therefore, $CaCl_2$ is simply calcium chloride, not calcium dichloride. As noted earlier in this chapter, prefixes are added in names of salts that contain more than 1 kind of cation or more than 1 kind of anion to show the relative numbers of ions. As an example, KH_2PO_4 is called potassium dihydrogen phosphate.

This page titled 2.5.4.10: Acids, Bases and Salts is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **4.10: Acids, Bases and Salts** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.4.11: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. What distinguishes the molecules of chemical compounds from those of elements, such as N_2?

2. Several "characteristics of compounds that meet the criteria of being green" were mentioned at the beginning of this chapter. Near the end of the chapter, acetic acid was mentioned as a "green acid." In what respects does it meet the criteria of green compounds?

3. What is sodium stearate? Why is it regarded as being green?

4. Which of the following is **not** usually regarded as a characteristic of green chemical compounds? Why is it not so regarded?

A. Preparation from renewable resources

B. Low tendency to undergo sudden, violent, unpredictable reactions

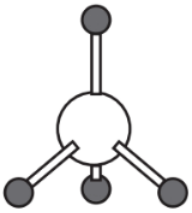
C. Readily biodegradable

D. Extremely high stability

5. What are valence electrons? Why are they particularly important?

6. What is the octet rule? Why is it particularly important in chemistry?

7. What does the structure representing CH_4 below say about bonding and octets of electrons around the central C atom?



8. Considering that the central nitrogen atom in ammonia, NH_3, has an unshared pair of valence electrons and 3 pairs shared between N and H, propose a structure for the ammonia molecule based upon the structure of the methane molecule in the preceding question. Use a pair of dots to represent the unshared pair of electrons.

9. What is an ionic bond? Why is it not regarded as being between one specific cation and a specific anion in an ionic compound?

10. Do ionic compounds such as NaCl obey the octet rule? Explain.

11. Why is NaCl referred to as a formula unit of the ionic compound rather than a molecule of sodium chloride?

12. Energy is involved in several steps of the process by which an elemental metal and an elemental nonmetal are converted to an ionic compound (salt). Of these, which has the largest energy?

13. Place the following ions in decreasing order of size: Na^+, Cl^-, Al^{3+}, K^+

14. What is a major disadvantage of calcium chloride as a road de-icing agent? Why is calcium acetate a good substitute?

15. List some important characteristics of a covalent bond.

16. What is the major characteristic of ions in ionic liquids that enable these materials to be liquid at around room temperature?

17. Can the atoms in NO_2 obey the octet rule? Suggest the structural formula for this molecule inwhich the 2 O atoms are bonded to an N atom.

18. Coordinate covalent bonds are normally regarded as those in which each of two atoms contributes electrons to be shared in the bond. Are there any circumstances in which this is not true? If so, give an example.

19. What are three major ways in which covalent bonds are characterized?

20. What are some of the ways in which the characteristics of covalent bonds are related to green chemistry?

21. Why are elements in the middle of periods of the periodic table less likely to form ionic compounds and more likely to form covalent compounds than those near either end of each period?

22. Predict the formula of the compound formed when H reacts with P and explain.

23. Although hydrogen chloride, HCl, exists as a gas, the contest for the two shared electrons in the bond between H and Cl is unequal, with the Cl nucleus having the greater attraction. Suggest the nature of the H-Cl bond and suggest what may happen when HCl gas dissolves in water to produce hydrochloric acid.

24. Using Lewis formulas, show the bonding in the SO_2 molecule in which two O atoms are bonded to a central O atom. Can another equivalent structure be drawn? Considering the bonding in NO discussed in this chapter, what are these structures called.

25. Summarize the information shown in the formula $Ca_3(PO_4)_2$.

This page titled 2.5.4.11: Questions and Problems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **Questions and Problems** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

2.5.5: Chemical Reactions- Making Materials Safely and Sustainable

“The materials that we make and the ways that we make them have an enormous impact on Earth’s environment. Much of green chemistry has to do with making materials safely and sustainably.”

2.5.5.1: Describing What Happens with Chemical Equations

2.5.5.2: Balancing Chemical Equations

2.5.5.3: Just Because You Can Write It Doesn't Mean That It Will Happen

2.5.5.4: Yield And Atom Economy in Chemical Reactions

2.5.5.5: Catalysts That Make Reactions Go

2.5.5.6: Kinds of Chemical Reactions

2.5.5.7: Oxidation-Reduction Reactions and Green Chemistry

2.5.5.8: Quantitative Information from Chemical Reactions

2.5.5.9: Energy in Chemical Reactions

2.5.5.10: Stoichiometry by the Mole Ratio Method

2.5.5.11: Limiting Reactant and Percent Yield

2.5.5.12: Titrations - Measuring Moles by Volume of Solution

2.5.5.13: Industrial Chemical Reactions - The Solvay Process

2.5.5.14: Questions and Problems

This page titled 2.5.5: Chemical Reactions- Making Materials Safely and Sustainable is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.5.1: Describing What Happens with Chemical Equations

How far would you have to go to find a diverse chemical factory carrying out hundreds of complex chemical processes? Not far, because your own body is just such a remarkably sophisticated factory that could not be duplicated by the efforts of thousands of chemists and chemical engineers and the expenditure of billions of dollars. As an example of a process that our bodies carry out consider the utilization of glucose sugar, chemical formula $C_6H_{12}O_6$, which is present in our blood and generates energy that the human body uses by the following metabolic biochemical reaction:

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$$

This is a **chemicalequation** that represents a chemical reaction, something that actually occurs with chemicals. It states that glucose reacts with molecular oxygen to produce carbon dioxide and water. The chemical reaction also produces energy and that is why the body carries it out to obtain the energy needed to move, work, and grow. The production of energy is sometimes denoted in the equation by adding "+ energy" to the right side.

Just as a chemical formula contains a lot of information about a chemical compound, a chemical equation contains much information about a chemical process. A chemical equation is divided into two parts by the arrow, which is read "yields." On the left of the arrow are the **reactants** and on the right are the **products**. A key aspect of a correctly written chemical equation is that it is **balanced**, with the same number of atoms of each element on the left as on the right. Consider the chemical equation above. The single molecule of $C_6H_{12}O_6$ contains 6 C atoms, 12 H atoms, and 6 O atoms. The 6 O_2 molecules contain 12 O atoms, giving a total of 18 O atoms among the reactants. Adding up all the atoms on the left gives 6 C atoms, 12 H atoms, and 18 O atoms among the reactants. On the right, the products contain 6 C atoms in the 6 CO_2 molecules, 12H atoms in the 6 H_2O molecules, and 12 O atoms in the 6 CO_2 molecules, as well as 6 O atoms in the 6 H_2O molecules, a total of 18 O atoms. So there are 6 C atoms, 12 H atoms, and 18 O atoms among the products, the same as in the reactants. Therefore, the equation is balanced.

An important exercise is the process of balancing a chemical equation. This consists of putting the correct numbers before each of the reactants and products so that equal numbers of each kind of atom are on both the left and right sides of the equation. The procedure for balancing a chemical equation is addressed in section 5.2.

Learning chemistry is largely an exercise in learning chemical language. In the chemical language the symbols of the elements are the alphabet. The formulas of the compounds are the words. And chemical equations are the sentences that tell what actually happens.

It is often important to know the physical states of reactants and products in chemical reactions. Suppose, for example, that a geologist tested a sample of rock to see if it were limestone by adding some liquid hydrochloric acid to the rock and observing the CO_2 gas coming off. The equation for the chemical reaction that occurred is

$$CaCO_3(s) + 2HCl(aq) \rightarrow CO_2(g) + CaCl_2(aq) + H_2O(l)$$

$$CaCO_3\,(s) + 2\,HCl\,(aq) \rightarrow CO_2\,(g) + CaCl_2\,(aq) + H_2O\,(l)$$

Here abbreviations in parentheses are used to represent the physical state of each reaction participant — (*s*) for solid, (*aq*) for a substance in solution, (*g*) for gas, and (*l*) for liquid. The equation above states that solid calcium carbonate reacts with an aqueous solution of hydrochloric acid to produce carbon dioxide gas, a solution of calcium chloride, and liquid water.

Chemical reactions often are **reversible**, that is, they may go either forward or backward. A reversible reaction is shown with a doublearrow, $\leftarrow \rightarrow$. Asan example, consider the reaction of dissolved ammonia, NH_3, with water to produce ammonium ion, $NH4^+$, and hydroxide ion, OH^-.

$$NH_3\,(aq) + H_2O\,(l) \rightleftharpoons NH_4{}^+\,(aq) + OH^-\,(aq)$$

Actually, only a small fraction of NH_3 molecules undergo this reaction at any given time, and those that are converted to $NH_4{}^+$ are rapidly converted back to NH_3. The double arrow in the chemical equation shows that both the forward and reverse processes occur. Another symbol that is sometimes used in chemical equations is Δ. This symbol denotes that heat is applied to make the

chemical reaction occur at a more rapid pace. It is normally placed over the arrow in the chemical reaction. Chemical equations are used to calculate the quantities of chemicals involved in a chemical reaction, either as reactants or as products. This is an important area of chemistry that is addressed by the topic of stoichiometry discussed later in this chapter in Section 5.8.

This page titled 2.5.5.1: Describing What Happens with Chemical Equations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.1: Describing What Happens with Chemical Equations** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.2: Balancing Chemical Equations

As noted earlier, a *balanced* chemical equation shows the same number of each kind of atom on both sides of the equation. The process of balancing chemical equations is an important exercise in chemistry and is addressed here.

Consider a simple example of balancing a chemical equation, the reaction of methane, CH_4, with elemental chlorine, Cl_2, to produce dichloromethane, CH_2Cl_2, an important laboratory solvent, and byproduct hydrogen chloride, HCl. The unbalanced chemical equation is

$$CH_4 + Cl_2 \rightarrow CH_2Cl_2 + HCl$$

Inspection of this equation as it is written shows that it is not balanced because it has 4 H on the left, but just 3 on the right and 2 Cl on the left, but 3 Cl on the right. In order to balance such an equation, consider one element at a time. Carbon is already balanced, so it is best to avoid changing any of the numbers in front of the C-containing compounds. The equation can be balanced for H by putting a 2 in front of HCl:

$$CH_4 + Cl_2 \rightarrow CH_2Cl_2 + 2\,HCl$$

Now everything is balanced except for Cl, of which there are 4 on the right, but just 2 on the left.Placing a 2 in front of Cl, gives the required 4 Cls on the left:

$$CH_4 + 2\,Cl_2 \rightarrow CH_2Cl_2 + 2\,HCl$$

This equation is now balanced with 1 C, 4 Hs, and 4 Cls on both the left and the right.

A crucial thing to remember in balancing a chemical equation is that *the chemical formulas must not be altered. Only the relative numbers of reactant and product species may be changed.*

Next consider the reaction of methane, CH_4, with iron oxide, Fe_2O_3, to give iron metal, Fe, carbon dioxide, CO_2, and water, H_2O. The unbalanced equation is

$$CH_4 + Fe_2O_3 \rightarrow Fe + CO_2 + H_2O$$

In this case it is helpful to note that CH_4 is the only source of both C and H and that 4 times as many H atoms as C atoms must appear in the products. That means that for each CO_2 there must be 2 H_2Os. Both C and H are balanced in the following:

$$CH_4 + Fe_2O_3 \rightarrow Fe + CO_2 + 2\,H_2O$$

But now O is not balanced. Furthermore, the 3 Os in Fe_2O_3 means that the number of O atoms must be divisible by 3, so try multiplying the three species balanced so far — CH_4, CO_2, and $2H_2O$ — by 3:

$$3\,CH_4 + Fe_2O_3 \rightarrow Fe + 3\,CO_2 + 6\,H_2O$$

That gives a total of 12 O atoms on the right, 6 each in 3 CO_2 and 6 H_2O. Taking 4 times Fe_2O_3 gives 12 Os on the left:

$$3\,CH_4 + 4\,Fe_2O_3 \rightarrow Fe + 3\,CO_2 + 6\,H_2O$$

The only species remaining to be balanced is Fe, which can be balanced by putting 8 in front of Fe on the right. The balanced equation is

$$3\,CH_4 + 4\,Fe_2O_3 \rightarrow 8\,Fe + 3\,CO_2 + 6\,H_2O$$

Checking the answer shows on both left and right 3 C, 8 Fe, 12 H, and 12 O demonstrating that the equation is in fact balanced.

Exercise: Balance the following:

1. $Fe_2O_3 + CO \rightarrow Fe + CO_2$
2. $FeSO_4 + O_2 + H_2O \rightarrow Fe(OH)_3 + H_2SO_4$
3. $C_2H_2 + O_2 \rightarrow CO_2 + H_2O$
4. $Mg_3N_2 + M_2O \rightarrow Mg(OH)_2 + NH_3$
5. $NaAlH_4 + H_2O \rightarrow H_2 + NaOH + Al(OH)_3$

6. $Zn(C_2H_5)_2 + O_2 \rightarrow ZnO + CO_2 + H_2O$

Answers: (1) $Fe_2O_3 + 3\,CO \rightarrow 2\,Fe + 3\,CO_2$, (2) $4\,FeSO_4 + O_2 + 10\,H_2O \rightarrow 4\,Fe(OH)_3 + 4\,H_2SO_4$, (3) $2\,C_2H_2 + 5\,O_2 \rightarrow 4\,CO_2 + 2\,H_2O$, (4) $Mg_3N_2 + 6\,H_2O \rightarrow 3\,Mg(OH)_2 + 2\,NH_3$ (5) $NaAlH_4 + 4\,H_2O \rightarrow 4\,H_2 + NaOH + Al(OH)_3$, (6) $Zn(C_2H_5)_2 + 7\,O_2 \rightarrow ZnO + 4\,CO_2 + 5\,H_2O$

This page titled 2.5.5.2: Balancing Chemical Equations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.2: Balancing Chemical Equations** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.3: Just Because You Can Write It Doesn't Mean That It Will Happen

The fact that a balanced chemical equation can be written does not necessarily mean that the chemical reaction that it represents will occur. As an example, it is known that a number of metals will react with acid to release elemental hydrogen gas and produce a metal salt. For example, if iron wire, Fe, is placed into a solution of sulfuric acid, H_2SO_4, H_2 gas is evolved,

$$\mathrm{Fe\,(s) + H_2SO_4\,(aq)\ \rightarrow H_2\,(g) + FeSO_4\,(aq)}$$

leaving $FeSO_4$ salt in solution. The copper salt, $CuSO_4$, is also known to exist. So one might believe that it could be prepared by reacting copper metal with H_2SO_4:

$$\mathrm{Cu\,(s) + H_2SO_4\,(aq)\ \rightarrow H_2\,(g) + CuSO_4\,(aq)}$$

This equation is balanced and it looks like it could occur. But, placing copper metal into a solution of H_2SO_4 in the laboratory results in — nothing. The reaction simply does not occur. The lesson here is that a balanced chemical equation is not sufficient reason to conclude that a reaction will take place.

Since $CuSO_4$ is known to exist, there has to be a way to prepare it. There are, in fact several ways. One pathway to the preparation of this salt starting with copper metal is to first react the copper with oxygen at a relatively high temperature to produce copper oxide:

$$\mathrm{2\,Cu\,(s) + O_2\,(g)\ \rightarrow 2\,CuO\,(s)}$$

The CuO product reacts with sulfuric acid to give $CuSO_4$ salt:

$$\mathrm{CuO\,(s) + H_2SO_4\,(aq)\ \rightarrow CuSO_4\,(aq) + H_2O\,(l)}$$

Alternate Reaction Pathways in Green Chemistry

Much of the science of green chemistry involves making decisions about alternative chemical reactions to choose a reaction or reaction sequence that provides maximum safety, produces minimum byproduct, and utilizes readily available materials. Consider two ways of preparing iron sulfate, $FeSO_4$. This chemical is commonly used to treat (clarify) water because when it is added to water and air is bubbled through the water, it produces $Fe(OH)_3$, a gelatinous solid that settles in the water and carries suspended mud and other particles with it. Consider two possible ways of making $FeSO_4$. The first of these was shown earlier and consists of the reaction of iron metal with sulfuric acid:

$$\mathrm{Fe\,(s) + H_2SO_4\,(aq)\ \rightarrow H_2\,(g) + FeSO_4\,(aq)}$$

leaving $FeSO_4$salt in solution. The copper salt, $CuSO_4$, is also known to exist. So one might believe that it could be prepared by reacting copper metal with H_2SO_4:

$$\mathrm{Cu\,(s) + H_2SO_4\,(aq)\ \rightarrow H_2\,(g) + CuSO_4\,(aq)}$$

This equation is balanced and it looks like it could occur. But, placing copper metal into a solution of H_2SO_4 in the laboratory results in — nothing. The reaction simply does not occur. The lesson here is that a balanced chemical equation is not sufficient reason to conclude that a reaction will take place.

Since $CuSO_4$ is known to exist, there has to be a way to prepare it. There are, in fact several ways. One pathway to the preparation of this salt starting with copper metal is to first react the copper with oxygen at a relatively high temperature to produce copper oxide:

$$\mathrm{2\,Cu\,(s) + O_2\,(g)\ \rightarrow 2\,CuO\,(s)}$$

The CuO product reacts with sulfuric acid to give $CuSO_4$ salt:

$$\mathrm{CuO\,(s) + H_2SO_4\,(aq)\ \rightarrow CuSO_4\,(aq) + H_2O\,(l)}$$

Alternate Reaction Pathways in Green Chemistry

Much of the science of green chemistry involves making decisions about alternative chemical reactions to choose a reaction or reaction sequence that provides maximum safety, produces minimum byproduct, and utilizes readily available materials. Consider

two ways of preparing iron sulfate, $FeSO_4$. This chemical is commonly used to treat (clarify) water because when it is added to water and air is bubbled through the water, it produces $Fe(OH)_3$, a gelatinous solid that settles in the water and carries suspended mud and other particles with it. Consider two possible ways of making $FeSO_4$. The first of these was shown earlier and consists of the reaction of iron metal with sulfuric acid:

$$\mathrm{Fe\,(s) + H_2SO_4\,(aq)\ \rightarrow H_2\,(g) + FeSO_4\,(aq)}$$

A second pathway would be to react iron oxide, FeO, with sulfuric acid:

$$\mathrm{FeO\,(s) + H_2SO_4\,(aq)\ \rightarrow FeSO_4\,(aq) + H_2O\,(aq)}$$

Which of these reactions would be the better choice? Both would work. The first reaction generates elemental H_2 gas as a byproduct. That has a potential downside because elemental hydrogen is highly explosive and flammable and could cause an explosion or fire hazard. But, in a contained reaction vessel that allowed for capture of H_2, the elemental hydrogen could be put to use as a fuel or reacted directly in a fuel cell to produce electricity (Section 3.2 and Figure 3.2). Furthermore, scrap iron metal and waste sulfuric acid are common materials that should be recycled and the synthesis of $FeSO_4$ by the direct reaction of the two can prepare a useful material from the two recyclable substances.

The second reaction (5.3.6) also gives the desired product. Its only byproduct is innocuous water. And there is no hazard from elemental hydrogen. In principle, the FeO required could be made by reacting scrap iron metal with oxygen from the air.

$$\mathrm{2\,Fe + O_2\ \rightarrow 2\,FeO}$$

but in practice the reaction tends to produce other oxides of iron, particularly Fe_2O_3 and Fe_3O_4.

This page titled 2.5.5.3: Just Because You Can Write It Doesn't Mean That It Will Happen is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.3: Just Because You Can Write It Doesn't Mean That It Will Happen** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.4: Yield And Atom Economy in Chemical Reactions

A fundamental concept basic to green chemistry that can be illustrated by chemical reactions is the distinction between yield and atom economy. In Chapter 2 *yield* was defined as a percentage of the degree to which a chemical reaction or synthesis goes to completion and *atomeconomy* was defined as the fraction of reactants that go into final products. Those two ideas are illustrated here for the preparation of HCl gas which, dissolved in water, produces hydrochloric acid. There are several ways in which HCl can be prepared. One of these commonly used in the laboratory is the reaction of concentrated sulfuric acid, H_2SO_4, with common table salt, NaCl, accompanied by heating to drive off the volatile HCl vapor:

$$2\,NaCl\,(s) + H_2SO_4\,(l) \rightarrow 2\,HCl\,(g) + Na_2SO_4\,(s)$$

This reaction can be performed so that all of the NaCl and H_2SO_4 react, which gives a 100% yield. But it produces Na_2SO_4 byproduct, so the atom economy is less than 100%. The percent atom economy is calculated very simply by the relationship

$$\text{Percent atom economy} = \frac{\text{Mass of desired product}}{\text{Total mass of product}} \times 100$$

(We could just as well divide by the total mass of reactants since in a chemical reaction it is equal to the total mass of products.) In this case, the mass of the desired product is that of 2 HCl and the total mass of product is that of 2HCl + Na_2SO_4. Given the atomic masses H 1.0, Cl 35.5, Na 23.0, and O 16.0 gives the following:

$$\text{Mass of desired product} = 2 \times (1.0 + 35.5) = 73.0$$

$$\text{Total mass product} = 2 \times (1.0 + 35.5) + (2 \times 23.0 + 32.0 + 4 \times 16.0) = 215$$

$$\text{Percent atom economy} = \frac{73.0}{215} \times 100 = 34.0\%$$

This result shows that even with 100% yield, the reaction is only 34.0% atom economical and if it were used as a means to prepare HCl large quantities of Na_2SO_4, a material with only limited value, would be produced. In contrast, the direct reaction of hydrogen gas with chlorine gas to give HCl gas,

$$H_2\,(g) + Cl_2\,(g) \rightarrow 2\,HCl\,(g)$$

can be carried out with 100% atom economy if all of the H_2 reacts with Cl_2. There is no waste byproduct.

This page titled 2.5.5.4: Yield And Atom Economy in Chemical Reactions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.4: Yield And Atom Economy in Chemical Reactions** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.5: Catalysts That Make Reactions Go

Carbon monoxide will certainly burn in the presence of oxygen from air as shown by the reaction

$$2\,CO + O_2 \rightarrow 2\,CO_2$$

Carbon monoxide is a product of automobile exhausts and an undesirable, toxic air pollutant. One way of ridding automobile exhaust gases of this pollutant is to pump air into the exhaust and convert the carbon monoxide to carbon dioxide as shown by the reaction above. However, even in the presence of oxygen, this reaction does not proceed to completion in an ordinary automobile exhaust system. It is enabled to occur, however, by passing the exhaust mixed with air over a solid honeycomb-like surface of ceramic coated with a metal that enables the reaction to occur, but is not itself consumed in the reaction. Such a substance is called a **catalyst**. Most people who have an automobile are vaguely aware that they have an automotive exhaust catalyst. They become much more acutely aware of this fact if the automobile's exhaust system fails an emissions test and the catalytic converter in it has to be replaced at a cost of several hundred dollars!

We do not have to go any farther than our own bodies to find catalysts. That is because all living organisms have biological catalysts that enable reactions to occur. Such living catalysts consist of specialized proteins called **enzymes**. Enzymes are discussed in Chapter 7.

A common enzyme-catalyzed process is the reaction of glucose (blood sugar, $C_6H_{12}O_6$) with molecular oxygen to produce energy mentioned at the beginning of this chapter:

This is the important process of **oxicrespiration** carried out by all organisms that live in contact with air and utilize oxygen from air to react with food materials. Although the overall reaction for oxic respiration can be written very simply, the actual process requires many steps and several catalytic enzymes are used. Other enzymes are used for various life processes, such as protein synthesis. There are enzymes that detoxify toxic substances, and in some cases they inadvertantly make toxic substances out of nontoxic ones. Some of the more common cancer-causing substances are actually synthesized from other molecules by enzyme action. Obviously enzymes are very important in life processes.

Catalysts speed up reactions. Depending upon the conditions the **rate of reaction** can vary significantly. Rates of chemical reactions are addressed by the area of **chemical kinetics**.

Catalysts are very important in green chemistry. One reason that this is so is because catalysts enable reactions to be carried out very specifically. Also, the right catalyst can enable reactions to occur with relatively less energy consumption and at relatively lower temperatures.

This page titled 2.5.5.5: Catalysts That Make Reactions Go is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.5: Catalysts That Make Reactions Go** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.6: Kinds of Chemical Reactions

It is useful to place chemical reactions in various categories. The important categories of chemical reactions are addressed here.

The simplest kind of chemical reaction to visualize is a **combinationreaction** in which two substances come together to form a new substance. The substances may be two elements, two compounds, or an element and a compound. An example of a combination reaction occurs when elemental carbon burns,

$$\mathrm{C} + \mathrm{O_2} \rightarrow \mathrm{CO_2}$$

to produce carbon dioxide. Since this reaction generates only one product, it occurs with 100% atom economy. Another combination reaction occurs when calcium oxide, CaO, present in a bed of solid material in a fluidized bed furnace used to burn coal reacts with sulfur dioxide:

$$\mathrm{CaO} + \mathrm{SO_2} \rightarrow \mathrm{CaSO_3}$$

The sulfur dioxide is a potential air pollutant produced from the burning of sulfur present in the coal. By injecting pulverized coal into a bed of CaO and other minerals kept in a fluid-like state by the injection of air, the sulfur dioxide produced has the opportunity to react with CaO and is not emitted as a pollutant with the stack gas.

In addition to being a combination reaction, the reaction above could also be called an **additionreaction** because the SO_2 adds to the CaO. Addition reactions are very desirable in the practice of green chemistry because they are 100% atom economical.

The opposite of a combination reaction is a **decomposition reaction**. An example of such a reaction occurs when a direct electrical current is passed between two electrodes through water to which a salt such as Na_2SO_4 has been added to make a solution in the water that is electrically conducting:

$$2\,\mathrm{H_2O\,(l)} \xrightarrow{Electrolysis} 2\,\mathrm{H_2\,(g)} + \mathrm{O_2\,(g)}$$

Reactions such as this that occur by the action of electricity passed through a solution are called **electrolysis** reactions. As written, the reaction is 100% atom economical. However, some side reactions may occur that reduce the efficiency. For example, impurity chloride ion, Cl^-, must be avoided in solution because it can produce some Cl_2 gas, a toxic, undesirable byproduct. Another inefficiency occurs because not all of the electricity passed through the solution is utilized to decompose water.

An example of a useful decomposition reaction is the high-temperature decomposition of methane,

$$\mathrm{CH_4\,(g)} \xrightarrow{\Delta} \mathrm{C\,(s)} + 2\,\mathrm{H_2\,(g)}$$

to produce elemental C and H_2 gas (where the triangle over the arrow shows that heat is applied —in this case to a temperature of 1260–1425°C — to make the reaction occur). The elemental carbon from this reaction is generated as a fine powder called **carbon black**. Carbon black is an ingredient of the paste in dry cells (such as those used in portable electronic devices); it is used as a filler in tires and to make electrodes for electrolysis processes such as the one by which aluminum metal is prepared.

Decomposition reactions do not always produce elements. For example, sodium bicarbonate mineral, $NaHCO_3$ may be heated,

$$2\,\mathrm{NaHCO_3\,(s)} \xrightarrow{\Delta} \mathrm{Na_2CO_3\,(s)} + \mathrm{CO_2\,(g)} + \mathrm{H_2O\,(g)}$$

to produce sodium carbonate, Na_2CO_3, commonly used as an industrial chemical to treat water, in cleaning solutions, and as an ingredient of glass.

> **Example:** Using atomic masses Na 23.0, H 1.0, C 12.0, and O 16.0, calculate the percent atom economy of the above reaction for the production of Na_2CO_3.
>
> *Answer:* When 2 formula units of $NaHCO_3$ react, 1 formula unit of Na_2CO_3 is produced. The masses involved in atomic mass units, u, are the following:
>
> $\text{Mass } 2\mathrm{NaHCO_3} = 2 \times (23.0 + 1.0 + 12.0 + 3 \times 16.0) = 168 \text{ u}$
>
> $Na_2CO_3 = 2 \times 23.0 + 12.0 + 3 \times 16.0 = 106 \text{ u}$
>
> $\text{Percent atom economy} = \frac{160 \text{ u}}{168 \text{ u}} \times 100 = 63.1\%$

A **substitution** or **replacement** reaction is one such as the reaction of iron and sulfuric acid,

$$\mathrm{Fe\,(s) + H_2SO_4\,(aq) \rightarrow H_2\,(g) + FeSO_4\,(aq)}$$

in which Fe replaces H in H_2SO_4, a reaction shown earlier for the preparation of $FeSO_4$. This reaction also falls under the classification of reactions involving **evolution of a gas**, in this case evolution of hydrogen gas. A **doublereplacement** reaction, also called a **metathesis** reaction, is one in which two compounds trade ions or other groups. When dissolved calcium chloride reacts with dissolved sodium carbonate,

$$\mathrm{CaCl_2\,(aq) + Na_2CO_3\,(aq) \rightarrow CaCO_3\,(s) + 2\,NaCl\,(aq)}$$

the Ca^{2+} ion in calcium chloride simply switches places with the Na^+ ions in the sodium carbonate to produce solid calcium carbonate and NaCl in solution. This is also a **precipitation** reaction in which a solid material forms from two substances dissolved in water; the solid formed is a **precipitate**. The removal of calcium from water as shown by this reaction is a common water treatment process called **water softening**. It is done because excessive levels of calcium cause formation of scale that can clog water pipes and damage plumbing apparatus. Whenever an acid and a base react, as shown here for the reaction of hydrochloric acid with sodium hydroxide,

$$\mathrm{HCl\,(aq) + NaOH\,(aq) \rightarrow H_2O\,(l) + NaCl\,(aq)}$$

water and a salt are formed. Such a reaction is a **neutralizationreaction** or simply an **acid-base** reaction.

> **Exercise:** Classify each of the following reactions as combination, decomposition, substitution, metathesis, neutralization, precipitation, or evolution of a gas. In some cases, a reaction will fit into more than one category.
>
> (a) $ce2Ca(s) + O2(g) \rightarrow 2CaO(s)$
>
> (b) $\mathrm{2\,KClO_3\,(s) \xrightarrow{\Delta} 2\,KCl\,(s) + 3\,O_2\,(g)}$
>
> (c) $\mathrm{SO_3\,(g) + H_2O\,(l) \rightarrow H_2SO_4\,(aq)}$
>
> (d) $\mathrm{MgCO_3\,(s) + 2\,HCl\,(aq) \rightarrow MgCl_2\,(aq) + H_2O\,(l) + CO_2\,(g)}$
>
> (e) $\mathrm{Zn\,(s) + CuCl_2\,(aq) \rightarrow Cu\,(s) + ZnCl_2\,(aq)}$
>
> (f) $\mathrm{KOH\,(aq) + Hcl\,(aq) \rightarrow KCl\,(aq) + H_2O\,(l)}$
>
> (g) $\mathrm{MgSO_4\,(aq) + 2\,KOH\,(aq) \rightarrow Mg(OH)_2\,(s) + K_2SO_4\,(aq)}$
>
> *Answers:* (a) Combination, (b) decomposition, evolution of a gas, (c) combination, (d)metathesis, evolution of a gas, (e) substitution, (f) neutralization, metathesis, (g) precipitation, metathesis

---

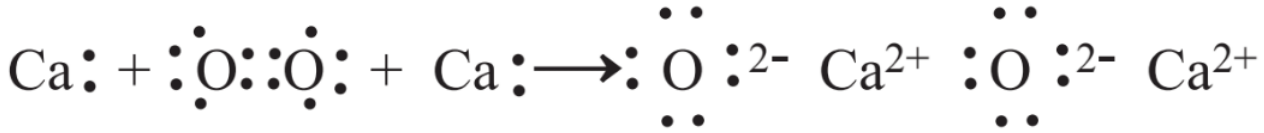
This page titled 2.5.5.6: Kinds of Chemical Reactions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.6: Kinds of Chemical Reactions** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.7: Oxidation-Reduction Reactions and Green Chemistry

Many reactions, including some of those given in the preceding section, are **oxidation-reduction** reactions, frequently called **redox** reactions. This name derives from the long standing use of *oxidation* to describe the reaction of a substance with oxygen. Consider the following reaction of elemental calcium with elemental oxygen:

$$2\,Ca + O_2 \rightarrow 2\,CaO$$



Combining with oxygen, Ca is *oxidized*. Whenever something is oxidized, something else has to be *reduced*. In this case, elemental oxygen is reduced to produce the oxide ion, O2-in CaO. It is seen from this reaction that the calcium atoms lose electrons when they are oxidized and the oxygen atoms gain electrons. This leads to another definition of oxidation-reduction reactions, which is that *when a chemical species loses electrons in a chemical reaction it is oxidized and when a species gains electrons it is reduced*.

Elemental hydrogen is commonly involved in oxidation-reduction. *Whenever a chemical species reacts with elemental hydrogen, it is reduced*. As an example, iron(II) oxide, FeO, can be reacted with elemental hydrogen,

$$FeO + H_2 \rightarrow Fe + H_2O$$

In this case the Fe in FeO is reduced to iron metal and the hydrogen in elemental H_2 is oxidized to H_2O.

When elemental oxygen reacts to produce chemically combined oxygen, it is acting as an **oxidizingagent** and is reduced. And when elemental hydrogen reacts to produce chemically combined hydrogen, it acts as a **reducingagent** and is oxidized. Consider what happens when the opposite reactions occur. When chemically combined oxygen is released as elemental oxygen from a chemical reaction, the oxygen is *oxidized*. And when elemental hydrogen is released as the result of a chemical reaction, hydrogen is *reduced*. A good illustration of these definitions may be seen when a direct electrical current is passed between two metal electrodes through water made electrically conducting by dissolving in it a salt, such as Na_2SO_4 as shown in Figure 5.1. At the left electrode, electrons are pumped into the system *reducing* the chemically bound H in H_2O to elemental H_2. An electrode at which *reduction* occurs is called the **cathode**. At the other electrode, electrons are removed from the system, elemental O_2 is released, and the oxygen in H_2O is *oxidized*. An electrode at which oxidation occurs is called the **anode.**

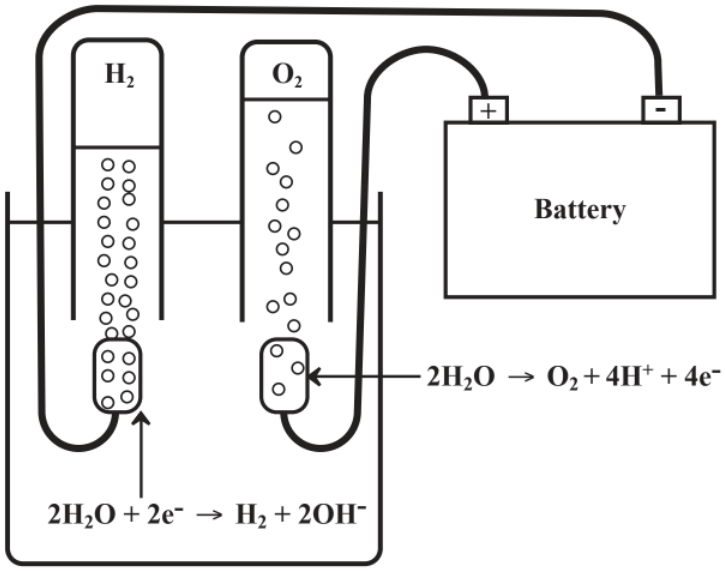


Figure 5.1. Electrolysis of water containing some dissolved salt to make it electrically conducting. At the left electrode(cathode) H in H_2O is reduced by adding electrons releasing H_2 gas. At the right electrode (anode) electrons are removed from chemically bound O in H_2O releasing elemental O_2 and the oxygen is oxidized.

The reaction shown above is an **electrolysis** reaction. It is very significant in the practice of green chemistry because it is a means of getting pure hydrogen and pure oxygen from water without the use of any other chemical reagents. For example, using a nonpolluting source of energy, such as wind power, elemental hydrogen can be generated for use in nonpolluting fuel cells (see Figure 3.2 and Chapter 16)

Oxidation-reduction reactions are very significant in energy conversion processes. An important example is photosynthesis,

$$6\,CO_2 + 6\,H_2O + h\,\nu \rightarrow C_6H_{12}O_6 + 6\,O_2$$

in which solar energy ($h\nu$) from sunlight is used by plants to produce glucose sugar, $C_6H_{12}O_6$, a high-energy compound that is used by organisms to provide energy for their metabolic needs. Since elemental oxygen is produced, oxygen is oxidized. Although it is not obvious based upon the discussion of oxidation-reduction so far, carbon is reduced; the carbon in the $C_6H_{12}O_6$ product is reduced compared to the carbon in the CO_2 reactant. The reverse of this reaction shown at the beginning of this chapter is

$$C_6H_{12}O_6 + 6\,O_2 \rightarrow 6\,CO_2 + 6\,H_2O + \text{energy}$$

which occurs when organisms—including humans—utilize glucose sugar to produce energy. In this case, oxygen reacts, an obvious oxidation process. The oxygen is reduced and carbon is oxidized by the action of the elemental oxygen.

A very common oxidation-reduction reaction occurs when fossil fuels are burned to produce energy. One such reaction occurs when natural gas (methane, CH_4) burns,

$$CH_4 + 2\,O_2 \rightarrow CO_2 + 2\,H_2O + \text{energy}$$

to produce carbon dioxide and water, releasing energy. The burning of gasoline, diesel fuel, coal,wood, and even hydrogen gas are oxidation-reduction reactions in which carbon or hydrogen are oxidized by the action of oxygen yielding usable energy.

Oxidation-reduction reactions are the most important kinds of reactions considered in green chemistry. That is true in part because of the central role currently played by the oxidation of fossil fuels and other materials in producing energy needed for chemical processes. Furthermore, the most common raw material currently used for making plastics, synthetic fabrics, and other manufactured materials is petroleum hydrocarbon. There are many hydrocarbon compounds all containing chemically bound carbon and hydrogen. A typical such compound is ethane, C_2H_6. The hydrogen and carbon in a hydrocarbon are in the most chemically reduced form, but required raw materials often are partially oxidized hydrocarbons in which O atoms are bonded to the hydrocarbon (complete oxidation of a hydrocarbon yields CO_2 and H_2O). Ethanol, C_2H_6O, used in chemical synthesis and as an oxygenated additive to make gasoline burn more smoothly with emission of fewer air pollutants is a partially oxidized hydrocarbon.

Large quantities of materials and energy are expended in converting petroleum hydrocarbons to partially oxidized compounds used as raw materials. For example, ethanol can be made from ethane taken from petroleum and natural gas by a series of chemical reactions for which the net process is the following:

$$2\,C_2H_6 + O_2 \rightarrow 2\,C_2H_6O$$

This transformation requires relatively severe conditions and a net loss of energy. A greener alternative is to use glucose sugar produced by photosynthesis (Reaction 5.7.3) to grow yeasts that produce an ethanol product

$$C_6H_{12}O_6 \rightarrow 2\,C_2H_6O + 2\,CO_2$$

a process that occurs under room temperature conditions. In addition to making ethanol, this fermentation process yields carbon dioxide in a concentrated form that can be used for carbonated beverages, supercritical carbon dioxide solvent, or pumped underground for tertiary petroleum recovery. The protein-rich yeast biomass produced in fermentation makes a good animal feed additive.

This page titled 2.5.5.7: Oxidation-Reduction Reactions and Green Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.7: Oxidation-Reduction Reactions and Green Chemistry** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.8: Quantitative Information from Chemical Reactions

Much of green chemistry is involved with calculations of quantities of materials involved in chemical reactions. It is essential to do such calculations in order to deal with the important concepts of percent yield and atom economy. Fortunately, it is easy to calculate quantities of materials if a balanced chemical reaction is known along with the pertinent atomic and formula masses. Both energy (Section 5.9) and mass (Section 5.10) in chemical reactions can be calculated.

To this point, we have been viewing chemical reactions in terms of individual atoms and molecules and have been thinking of masses in atomic mass units, u, used to express the masses of individual atoms and molecules. But that is much too small a scale to use in the laboratory. The chemist conveniently deals with grams and moles where a mole of a substance (see Section 4.8) typically has a mass of several to several hundred grams. Much of the remainder of this chapter deals with quantitative information from chemical reactions. Energy changes in chemical reactions are addressed first followed by consideration of masses of reactants and products.

This page titled 2.5.5.8: Quantitative Information from Chemical Reactions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.8: Quantitative Information from Chemical Reactions** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.9: Energy in Chemical Reactions

In addition to changes in the distribution of mass among various chemical species that occur with chemical reactions, another important participant in chemical reactions is **energy**. Chapter 15 defines and discusses energy in more detail. Here energy is considered in the form of **heat** manifested by the movement of atoms and molecules and as **chemical energy**, which is energy stored in chemical bonds in matter. The standard unit of energy is the **joule**, abbreviated **J**. A total of 4.184 J of heat energy will raise the temperature of 1 g of liquid water by 1°C.

As an example of energy involved with chemical reactions, Consider what happens in a burner on a kitchen range fueled by natural gas. The flame is obviously hot; something is going on that is releasing heat energy. The flame is also giving off light energy, probably as a light blue glow. A chemical reaction is taking place as the methane in the natural gas combines with oxygen in the air,

$$CH_4 + 2\,O_2 \rightarrow CO_2 + 2\,H_2O$$

to produce carbon dioxide and water. Most of the energy released during this chemical reaction is released as heat, and a little bit as light. It is reasonable to assume that the methane and oxygen contain stored energy as **chemical potential energy** and that it is released in producing carbon dioxide and water. Common sense tells us that it would be hard to get heat energy out of either of the products. They certainly won't burn! Water is used to put out fires, and carbon dioxide is even used in fire extinguishers.

The potential energy contained in chemical species is contained in the chemical bonds of the molecules that are involved in the chemical reaction. Figure 5.2 shows the kinds of bonds involved in methane, elemental oxygen, carbon dioxide, and water and the energy contained in each. The bond energies are in units of the number of kilojoules (thousands of joules, kJ) required to break a mole (6.02×1023, see Section 4.8) of the bonds (kJ/mol). The same amount of energy is released when a mole of a bond is formed. By convention, energy put into a system is given a positive sign and energy released is denoted by a negative sign.

To calculate the energy change when a mole of methane reacts with oxygen as shown in Reaction 5.9.1, the difference is taken between the sum of the energies of the bonds in the products and the sum of the energies of the bonds in the reactants. Examination of Reaction 5.9.1 and Figure 5.2 shows the following total bond energies in the products:

$$1\ \text{mol}\ CO_2 \times \frac{2\ \text{mol C=O}}{\text{mol}\ CO_2} \times \frac{799\ \text{kJ}}{\text{mol C=O}} = 1598\ \text{kJ}$$

$$2\ \text{mol}\ H_2O \times \frac{2\ \text{mol O-H}}{\text{mol}\ H_2O} \times \frac{459\ \text{kJ}}{\text{mol O-H}} = 1836\ \text{kJ}$$

A similar calculation gives the total bond energies in the reactants

$$1\ \text{mol}\ CH_4 \times \frac{4\ \text{mol C-H}}{1\ \text{mol}\ CH_4} \times \frac{411\ \text{kJ}}{\text{mol C-H}} = 1644\ \text{kJ}$$

$$2\ \text{mol}\ O_2 \times \frac{1\ \text{mol O=O}}{\text{mol}\ O_2} \times \frac{494\ \text{kJ}}{\text{mol O=O}} = 988\ \text{kJ}$$

Total bond energy in reactants = 1644 kJ + 988 kJ = 2632 kJ

The difference in bond energies between products and reactants is 3434 kJ - 2632 kJ = 802 kJ

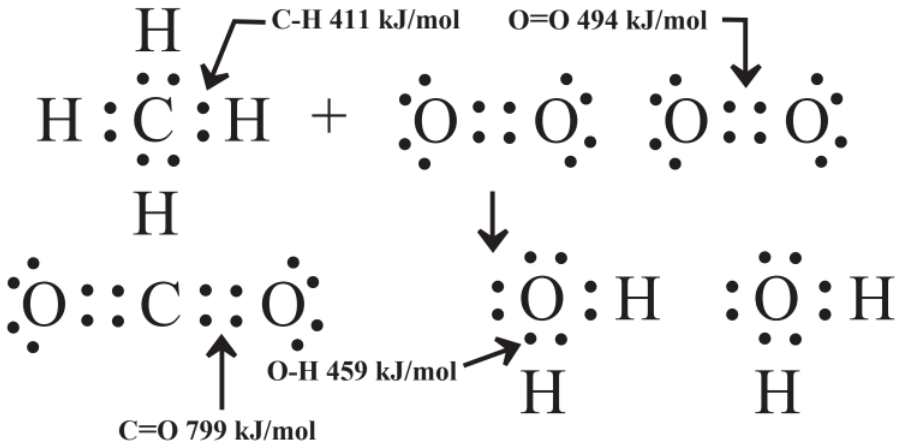


Figure 5.2. Bonds and their energies in the chemical species involved when methane burns in oxygen to produce carbon dioxide and water.

This calculation states that, based upon considerations of bond energy, alone, the energy released when 1 mole of CH_4 reacts with 2 moles of O_2 to produce 1 mole of CO_2 and 2 moles of H_2O, is 802 kJ. This is an **exothermic** reaction in which heat energy is *released*, so it is denoted as-802 kJ. This value is close to the value that would be obtained by experimentally measuring the heat energy released by the reaction, assuming all the reactants and products were in the gas phase. (A significant amount of heat energy is released when vapor-phase water condenses to liquid. Highly efficient condensing gas furnaces capture this heat in a heat exchanger where the water vapor in the furnace exhaust condenses to the liquid phase.) For the most part, therefore, the amount of heat energy released in a chemical reaction, and the amount of potential chemical energy contained in the reactants is equal to the difference between the total bond energies of the products and those of the reactants.

This page titled 2.5.5.9: Energy in Chemical Reactions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.9: Energy in Chemical Reactions** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.10: Stoichiometry by the Mole Ratio Method

The calculation of quantities of materials involved in chemical reactions is called **stoichiometry.** To illustrate stoichiometric calculations consider a typical chemical reaction, in this case the heat-producing combustion of ethane, a hydrocarbon fuel with a chemical formula of C_2H_6:

$$2\,C_2H_6 + 7\,O_2 \rightarrow 4\,CO_2 + 6\,H_2O$$

Rather than viewing this reaction in terms of individual molecules, it is possible to scale up to moles. Recall from Section 4.8 that the mole is a fundamental unit for quantity of material and that each mole contains Avogadro's number (6.022×10^{23}) of formula units (molecules of covalently bound compounds). This equation simply says that 2 moles of C_2H_6 react with 7 moles of O_2 to yield 4 moles of CO_2 and 6 moles of H_2O. Now we can examine the equation in more detail to do some quantitative calculations. Before doing that, however, review the following two terms

Formula mass: The sum of the atomic masses of all the atoms in a formula unit of a compound. Although the average masses of atoms and molecules may be expressed in atomic mass units (amu or u), formula mass is generally viewed as being relative and without units.

Molar mass: Where X is the formula mass, the molar mass is X grams of an element or compound, that is, the mass in grams of 1 mole of the element or compound

Given the atomic masses H 1.0, C 12.0, and O 16.0 the molar mass of C_2H_6 is $2\times12.0 + 6\times1.0 = 30.0$ g/mol, that of O_2 is $2\times16.0 = 32.0$ g/mol, that of CO_2 is $12.0 + 2\times16.0 = 44.0$ g/mol, and that of H_2O is $2\times1.0 + 16.0 = 18.0$ g/mol. Now consider the chemical equation

$$2\,C_2H_6 + 7\,O_2 \rightarrow 4\,CO_2 + 6\,H_2O$$

in terms of the minimum whole number of moles reacting and produced and the masses in grams of these quantities. The equation states that 2 moles of C_2H_6 with a mass of 2×30.0 g = 60.0 g of C_2H_6 react with 7 moles of O_2 with a mass of 7×32.0 g = 224 g of O_2 to produce 4 moles of CO_2 with a mass of 4×44.0 g = 176 g of CO_2 and 6 moles of H_2O with a mass of 6×18.0 g = 108 g of H_2O. The total mass of reactants is

$$60.0 \text{ g of } C_2H_6 + \ 224 \text{ g of } O_2 = 284.0\text{g of reactants}$$

and the total mass of products is

$$176 \text{ g of } CO_2 + 108\text{g of } H_2O = 284 \text{ g of products}$$

Note that, as in all chemical reactions, the total mass of products equals the total mass of reactants. Stoichiometry, the calculation of quantities of materials involved in chemical reactions, is based upon the *law of conservation of mass* which states that *the total mass of reactants in a chemical reaction equals the total mass of products*, because matter is neither created nor destroyed in chemical reactions. The basic premise of the **mole ratio method** of stoichiometric calculations is that the relative numbers of moles of reactants and products remain the same regardless of the total quantity of reaction. To illustrate stoichiometric calculations, consider again the following reaction:

$$2\,C_2H_6 + 7\,O_2 \rightarrow 4\,CO_2 + 6\,H_2O$$

As noted above, this equation states that 2 moles C_2H_6 react with 7 moles of O_2 to produce 4 moles of CO_2 and 6 moles of H_2O. The same ratios hold true regardless of how much material reacts. So for 10 times as much material, 20 moles C_2H_6 react with 70 moles of O_2 to produce 40 moles of CO_2 and 60 moles of H_2O

Suppose that it is given that 18.0 g of C_2H_6 react. What is the mass of O_2 that will react with this amount of C_2H_6? What mass of CO_2 is produced? What mass of H_2O is produced? This problem can be solved by the **mole ratio method**. Mole ratios are, as the name implies, simply the ratios of various moles of reactants and products to each other as shown by a chemical equation. Mole ratios are obtained by simply examining the chemical equation in question; the three that will be used in solving the problem posed are the following:

Suppose that it is given that 18.0 g of C_2H_6 react. What is the mass of O_2 that will react with this amount of C_2H_6? What mass of CO_2 is produced? What mass of H_2O is produced? This problem can be solved by the **mole ratio method**. Mole ratios are, as the

name implies, simply the ratios of various moles of reactants and products to each other as shown by a chemical equation. Mole ratios are obtained by simply examining the chemical equation in question; the three that will be used in solving the problem posed are the following:

$\frac{7 \text{mol } O_2}{2 \text{ mol } C_2H_6}$ $\frac{4 \text{ mol } CO_2}{2 \text{ mol } C_2H_6}$ $\frac{6 \text{ mol } H_2O}{2 \text{ mol } C_2H_6}$

To solve for the mass of O_2 reacting the following steps are involved:

A. Mass of C_2H_6 reacting B. Convert to moles of C_2H_6 C. Convert to moles of O_2 D. Convert to mass of O_2

In order to perform the calculation, it will be necessary to have the molar mass of C_2H_6, stated earlier as 30.0 g/mol, the molar mass of O_2(18.0 g/mol) and the mole ratio relating moles of O_2 reactant to moles of C_2H_6, 7 mol O_2/2 mol C_2H_6. The calculation becomes the following

$$\text{Mass of } O_2 = 18.0 \text{ g } C_2H_6 \times \frac{1 \text{ mol } C_2H_6}{30.0 \text{ g } C_2H_6} \times \frac{7 \text{ mol } O_2}{2 \text{ mol } C_2H_6} \times \frac{32.0 \text{ g } O_2}{1 \text{ mol } O_2} = 67.2 \text{ g } O_2$$

Note that in this calculation units cancel above and below the line, starting with units of g C_2H_6. Now that the mass of O_2 reacting has been calculated, it is possible using the appropriate mole ratios and molar masses to calculate the masses of CO_2 and of H_2O produced as follows:

$$\text{Mass of } CO_2 = 18.0 \text{ g } C_2H_6 \times \frac{1 \text{ mol } C_2H_6}{30.0 \text{ g } C_2H_6} \times \frac{6 \text{ mol } H_2O}{2 \text{ mol } C_2H_6} \times \frac{18.0 \text{ g } H_2O}{1 \text{ mol } H_2O} = 32.4 \text{ g } H_2O$$

$$\text{Mass of } H_2O = 18.0 \text{ g } C_2H_6 \times \frac{1 \text{ mol } C_2H_6}{30.0 \text{g} C_2H_6} \times \frac{6 \text{ mol } H_2O}{2 \text{ mol } C_2H_6} \times \frac{18.0 \text{ g } H_2O}{1 \text{ mol } H_2O} = 32.4 \text{ g } H_2O$$

Are the masses calculated above correct? A good check is to compare the total mass of reactants, 18.0 g C_2H_6 + 67.2 g O_2= 85.2 g of reactants, with the total mass of products, 52.8 g CO_2 + 32.4 g H_2O = 85.2 g of products. The fact that the total mass of reactants is equal to the total mass of products gives confidence that the calculations are correct.

As one more example consider the reaction of 15.0 g of Al with Cl_2 to give $AlCl_3$:

$$2\,Al + 3\,Cl_2 \rightarrow 2\,AlCl_3$$

What mass of Cl_2 reacts and what is the mass of $AlCl_3$ produced? The atomic mass of Al is 27.0 and that of Cl is 35.5. Therefore, the molar mass of Cl_2 is 71.0 g/mol and the molar mass of $AlCl_3$ is 133.5 g/mole. The mass of Cl_2 reacting is

$$\text{Mass of } Cl_2 = 15.0 \text{ g Al} \times \frac{1 \text{ mol Al}}{27.0 \text{ g Al}} \times \frac{3 \text{ mol } Cl_2}{2 \text{ mol Al}} \times \frac{71.0 \text{ g } Cl_2}{1 \text{ mol } Cl_2} = 59.2 \text{ g } Cl_2$$

$$\text{Mass of } AlCl_3 = 15.0 \text{ g Al} \times \frac{1 \text{ mol Al}}{27.0 \text{ g Al}} \times \frac{2 \text{ mol } AlCl_3}{2 \text{ mol Al}} \times \frac{133.5 \text{ g } AlCl_3}{1 \text{ mol } AlCl_3} = 74.2 \text{ g } AlCl_3$$

As a check, 15.0 g Al + 59.2 g Cl_2 reactant gives a total of 74.2 g of reactants equal to the mass of the $AlCl_3$ product.

Exercise

Calculate the mass of CH_4 that reacts and the masses of the products when 25.0 g of Fe_2O_3 undergo the reaction below. The atomic masses involved are H 1.0, C12.0, O 16.0, Fe 55.8

$$4\,Fe_2O_3 + 3\,CH_4 \rightarrow 8\,Fe + 3\,CO_2 + 6\,H_2O$$

Answer

Answer: 1.88 g CH4, 17.5 g Fe, 5.2 g CO2, 4.2 g H2O

Exercise

Calculate the mass of O_2 that reacts and the masses of the products when 100 g of benzoic acid, $C_7H_6O_2$ undergo the reaction below. The atomic masses involved are H 1.0, C 12.0, and O 16.0.

$$2\,C_7H_6O_2 + 15\,O_2 \rightarrow 14\,CO_2 + 6\,H_2O$$

Answer

197 g O_2, 252 g CO_2, 44.3 g H_2O

This page titled 2.5.5.10: Stoichiometry by the Mole Ratio Method is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.10: Stoichiometry by the Mole Ratio Method** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.11: Limiting Reactant and Percent Yield

Mixing of exact amounts of reactants such that all are consumed and none left over in a chemical reaction almost never occurs. Instead, one of the reactants is usually a **limiting reactant**. Suppose, for example that 100 g of elemental zinc (atomic mass 65.4) and 80 g of elemental sulfur (atomic mass 32.0) are mixed and heated undergoing the following reaction:

$$\mathrm{Zn + S \ \rightarrow ZnS}$$

What mass of ZnS, formula mass 97.4 g/mol, is produced? If 100 g of zinc react completely, the mass of S reacting and the mass of ZnS produced would be given by the following calculations:

$$\text{Mass s} = 100.0 \text{ G Zn} \times \frac{1 \text{ mol Zn}}{65.4 \text{ g Zn}} \times \frac{1 \text{ mol S}}{1 \text{ mol Zn}} \times \frac{32.0 \text{ g S}}{1 \text{ mol S}} = 48.9 \text{ g S}$$

$$\text{Mass ZnS} = 100.0\text{g Zn} \times \frac{1 \text{ mol Zn}}{65.4 \text{ g Zn}} \times \frac{1 \text{ mol S}}{1 \text{ mol Zn}} \times \frac{97.4 \text{ g ZnS}}{1 \text{ mol ZnS}} = 149 \text{ g S}$$

Only 48.9 g of the available S react, so sulfur is in excess and zinc is the limiting reactant. A similar calculation for the amount of Zn required to react with 80 g of sulfur would show that 164g of Zn would be required, but only 100 g is available.

Exercise

A solution containing 10.0 g of HCl dissolved in water (a solution of hydrochloric acid) was mixed with 8.0 g of Al metal undergoing the reaction

$$\mathrm{2\,Al + 6\,HCl \ \rightarrow 2\,AlCl_3 + 3\,H_2}$$

Given atomic masses H 1.0, Al 27.0, and Cl 35.5, which reactant was left over? How much? What mass of $\mathrm{AlCl_3}$ was produced?

Answer

HCl was the limiting reactant. Only 2.47 g of Al were consumed leaving 5.53 g of Al unreacted. The mass of $\mathrm{AlCl_3}$ produced was 12.2 g

Percent Yield

The mass of product calculated from the mass of limiting reactant in a chemical reaction is called the **stoichiometric yield** of a chemical reaction. By measuring the actual mass of a product produced in a chemical reaction and comparing it to the mass predicted from the stoichiometric yield it is possible to calculate the *percent yield*. This concept is illustrated by the following example.

Suppose that a water solution containing 25.0 g of $\mathrm{CaCl_2}$ was mixed with a solution of excess sodium sulfate,

$$\mathrm{CaCl_2\,(aq) + Na_2SO_4\,(aq) \ \rightarrow CaSO_4\,(s) + 2\,NaCl\,(aq)}$$

to produce a solid precipitate of $\mathrm{CaSO_4}$, the desired product of the reaction. (Recall that a *precipitate* is a solid formed by the reaction of species in solution; such a solid is said to *precipitate* from the solution.) Removed by filtration and dried, the precipitate was found to have a mass of 28.3 g, the **measured yield**. What was the percent yield?

Using atomic masses Ca 40.0, Cl 35.5, Na 23.0, and O, 16.0 gives molar masses of 111 g/mol for $\mathrm{CaCl_2}$ and 136 g/mol for $\mathrm{CaSO_4}$. Furthermore, 1 mole of $\mathrm{CaCl_2}$ yields 1 mol of $\mathrm{CaSO_4}$. The stoichiometric yield of $\mathrm{CaSO_4}$ is given by the following calculation

$$\text{Mass } \mathrm{CaSO_4} = 25.0 \text{ g } \mathrm{CaCl_2} \times \frac{1 \text{ mol } \mathrm{CaCl_2}}{111 \text{ g } \mathrm{CaCl_2}} \times \frac{1 \text{ mol } \mathrm{CaSO_4}}{1 \text{ mol } \mathrm{CaCl_2}} \times \frac{136 \text{ g } \mathrm{CaSO_4}}{1 \text{ mol } \mathrm{CaSO_4}} = 30.6 \text{ g } \mathrm{CaSO_4}$$

The percent yield is calculated by the following:

$$\text{Percent yield} = \frac{\text{Measured yield}}{\text{Stoichiometric yield}} \times 100$$

$$\text{Percent yield} = \frac{28.3 \text{ g}}{30.6 \text{ g}} \times 100 = 92.5\%$$

This page titled 2.5.5.11: Limiting Reactant and Percent Yield is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.11: Limiting Reactant and Percent Yield** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.12: Titrations - Measuring Moles by Volume of Solution

Masses are commonly measured with a laboratory balance that registers in grams. Masses of industrial chemicals are measured with much larger industrial scales that commonly give masses in kilograms or tons. In doing laboratory stoichiometric measurements with species in solution, it is often convenient to measure volumes of solution rather than masses of reactants. Solutions can be prepared that contain known numbers of moles per unit volume of solution. The volume of the reagent that must be added to another reagent to undergo a particular reaction can be measured with a device called a **buret**. A buret is shown in Figure 5.3. By measuring the volume of a solution of known concentration of solute required to react with another reactant, the number of moles of solute reacting can be calculated and stoichiometric calculations can be performed based upon the reaction. This procedure is commonly used in chemical analysis and is called **titration**.

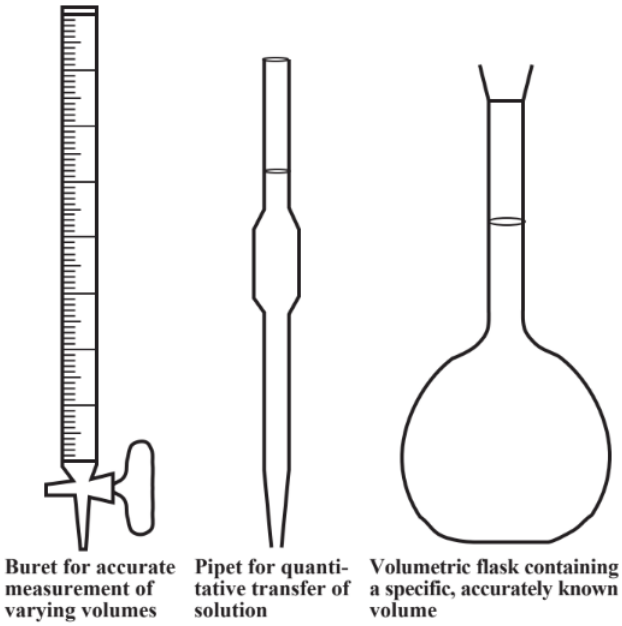


Figure 5.3. A buret consists of a narrow glass tube marked off in divisions of milliliters (mL) further subdivided into tenths of mL, usually with a total capacity of 50 mL. It enables accurate measurements of volumes of solution delivered through a stopcock estimated to the nearest 0.01 mL. A pipet delivers a fixed volume of solution. A volumetric flask contains a fixed volume of solution.

It is especially easy to relate volumes of solutions stoichiometrically when the solution concentrations are expressed as *molar concentration*, *M*. This concentration unit is defined as

$$\mathrm{M} = \frac{\text{moles of solute}}{\text{number of liters of solution}}$$

The number of moles of a substance, in this case the moles of solute, is related to the mass of the substance by

$$\text{Moles of solute} = \frac{\text{moles of solute}}{\text{molar mass of solute, g/mol}}$$

These two relationships can be combined to give the following useful equation:

$$\mathrm{M} = \frac{\text{mass of solute}}{(\text{molar mass of solute}) \times (\text{number of liters of solution})}$$

A solution of known concentration that is added to a reaction mixture during the procedure of titration is a **standard solution**. One of the most common of these is a standard base solution of sodium hydroxide, NaOH. Typically, the concentration of sodium hydroxide in such a standard solution is 0.100 mol/L. Suppose that it is desired to make exactly 2 liters of a solution of 0.100mol/L sodium hydroxide. What mass of NaOH, molar mass 40.0 g/mol, is dissolved in this solution? To do this calculation, use Equation 5.12.3 rearranged to solve for mass of solute:

$$\text{Mass NaOH} = \mathrm{M} \times (\text{molar mass NaOH}) \times (\text{liters NaOH})$$

$$\text{Mass NaOH} = 0.100 \text{ mol/L} \times 40.0 \text{ g/mol} \times 2.00 \text{ L} = 8.00 \text{ g NaOH}$$

A common titration procedure is to use a standard solution of base to titrate an unknown solution of acid or to use standard acid to determine base. As an example consider an analysis for acid of a sample of water used to scrub exhaust gas from a hospital incinerator. The water is acidic because of the presence of hydrochloric acid produced by the scrubbing of HCl gas from the incinerator stack gas where the HCl was produced in the burning of polyvinyl chloride in the incinerator. Suppose that a sample of 100 mL of the scrubber water was taken for titration with a 0.125 mol/L standard NaOH and that the volume of standard NaOH consumed was 11.7 mL. What was the molar concentration of HCl in the stack gas scrubber water? To solve this problem it is necessary to know that the reaction between NaOH and HCl is,

$$\text{NaOH} + \text{HCl} \rightarrow \text{NaCl} + \text{H}_2\text{O}$$

a neutralization reaction in which water and a salt, NaCl are produced. Examination of the reaction shows that 1 mole of HCl reacts for each mole of NaOH. Equation 5.12.1 applies to both the standard NaOH solution and the HCl solution being titrated leading to the following equations:

$$M_{HCl} = \frac{moles_{HCl}}{liters_{HCl}} = \text{and } \text{M}_{NaOH} = \frac{moles_{NaOH}}{liters_{NaOH}}$$

When exactly enough NaOH has been added to react with all the HCl present, the reaction is complete with no excess of either HCl or NaOH. In a titration this **endpoint** is normally shown by the change of color of a dye called an **indicator** dissolved in the solution being titrated. At the endpoint moles HCl= moles NaOH and the two equations above can be solved to give,

$$\text{M}_{\text{HCl}} \times \text{liters}_{\text{HCl}} = \text{M}_{\text{NaOH}} \times \text{liters}_{\text{NaOH}}$$

which can be used to give the molar concentration of HCl:

$$\text{M}_{\text{HCl}} = \frac{\text{M}_{\text{NaOH}} \times \text{liters}_{\text{NaOH}}}{\text{liters}_{\text{HCl}}}$$

Converting the volumes given from mL to liters and substituting into this equation gives the molar concentration of HCl in the incinerator scrubber water:

$$\text{M}_{\text{HCl}} = \frac{0.125\text{mol/L} \times 0.0117\text{L}}{0.100 \text{ L}} = 0.0146\text{mol/L}$$

Determining Percentage Composition by Titration

A useful application of titration, or **titrimetric analysis** as it is called, is to determine the percentage of a substance in a solid sample that will react with the titrant. To see how this is done, consider a sample consisting of basic lime, $Ca(OH)_2$, and dirt with a mass of 1.26 g. Using titration with a standard acid solution it is possible to determine the mass of basic $Ca(OH)_2$ in the sample and from that calculate the percentage of $Ca(OH)_2$ in the sample. Assume that the solid sample is placed in water and titrated with 0.112 mol/L standard HCl, a volume of 42.2 mL(0.0422 L) of the acid being required to reach the end point. The $Ca(OH)_2$ reacts with the HCl

$$\text{Ca(OH)}_2 + 2\,\text{HCl} \rightarrow \text{CaCl}_2 + 2\,\text{H}_2\text{O}$$

whereas the dirt does not react. Examination of this reaction shows that at the end point the mole ratio

$$\frac{1 \text{ mol Ca(OH)}_2}{2 \text{ mol HCl}}$$

applies. At the end point, the number of moles of HCl can be calculated from

$$\text{Mol}_{\text{HCl}} = \text{liters}_{\text{HCl}} \times \text{M}_{\text{HCl}}$$

and, since the molar mass of $Ca(OH)_2$ is 74.1 (given atomic masses 40.1, 16.0, and 1.0 for Ca, O, and H, respectively), the mass of $Ca(OH)_2$ is given by

$$\text{Mass}_{\text{Ca(OH)}_2} = \text{moles}_{\text{Ca(OH)}_2} \times \text{molar mass}_{\text{Ca(OH)}_2}$$

With this information it is now possible to calculate the mass of $Ca(OH)_2$:

$$\text{Mass}_{\text{Ca(OH)}_2} = \text{mol}_{\text{Ca(OH)}_2} \times \frac{74.1 \text{ g Ca(OH)}_2}{1 \text{ mol Ca(OH)}_2}$$

$$\text{Mass}_{\text{Ca(OH)}_2} = \underbrace{\text{Liters}_{\text{HCl}} \times \text{M}_{\text{HCl}}}_{\text{Moles HCl reacting}} \times \underbrace{\frac{1 \text{ mol Ca(OH)}_2}{2 \text{ mol HCl}}}_{\text{Converts from moles HCl to moles Ca(OH)}_2} \times \underbrace{\frac{74.1 \text{ g Ca(OH)}_2}{1 \text{ mol Ca(OH)}_2}}_{\text{Gives mass Ca(OH)}_2\text{from moles Ca(OH)}_2}$$

$$\text{Mass}_{\text{Ca(OH)}_2} = 0.0422\text{L HCl} \times \frac{0.112 \text{ mol HCl}}{1 \text{ L HCl}} \times \frac{1 \text{ mol Ca(OH)}_2}{2 \text{ mol HCl}} \times \frac{74.1 \text{ g Ca(OH)}_2}{1 \text{ mol Ca(OH)}_2}$$

$$\text{Mass Ca(OH)}_2 = 0.175\, g$$

$$\text{Percent}_{\text{Ca(OH)}_2} = \frac{\text{mass Ca(OH)}_2}{\text{mass sample}} \times \frac{0.175g}{1.26g} \times 100 = 13.9\%$$

Exercise

Exercise: A 0.638 g sample consisting of oxalic acid, H_2C_2O4, and sodium oxalate, $Na_2C_2O_4$ was dissolved and titrated with 0.116 mol/L sodium hydroxide, of which 47.6mL (0.0476 L) was required. Each molecule of $H_2C_2O_4$ releases 2 H^+ ions. Calculate the percentage of oxalic acid in the sample.

Answer

38.9%

This page titled 2.5.5.12: Titrations - Measuring Moles by Volume of Solution is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.12: Titrations - Measuring Moles by Volume of Solution** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.13: Industrial Chemical Reactions - The Solvay Process

Literally thousands of chemical reactions are used to make important industrial products. Most of these involve organic chemicals, which are addressed in Chapter 9 and later chapters of this book. Some are used to make inorganic chemicals in large quantities. One such synthesis operation is the **Solvay process**, long used to make sodium bicarbonate and sodium carbonate, industrial chemicals required for glass making, cleaning formulations, and many other applications. The Solvay process is examined in some detail in this section because it illustrates some important inorganic chemical reactions and can be used for the discussion of green chemistry in industry.

The key reaction in Solvay synthesis is,

$$NaCl + NH_3 + Co_2 + H_2O \rightarrow NaHCO_3\,(s) + NH_4Cl$$

in which a sodium chloride solution (brine) is saturated with ammonia gas (NH_3), then with carbon dioxide, and finally cooled. This is a precipitation reaction in which solid sodium bicarbonate, $NaHCO_3$, comes out of solution. When heated, the solid $NaHCO_3$ yields solid sodium carbonate, Na_2CO_3, water vapor, and carbon dioxide gas:

$$2\,NaHCO_3 + \text{heat} \rightarrow Na_2CO_3 + H_2O\,(g) + CO_2\,(g)$$

In keeping with the practice of green chemistry (although Solvay developed the process long before anyone ever thought of green chemistry), the CO_2 from Reaction 5.13.2 is recycled back into Reaction 5.13.1

The raw materials for the Solvay process are cheap. The NaCl solution can be pumped from the ground from brine deposits in some locations, or fresh water can be pumped into a salt formation to dissolve NaCl and the resulting brine pumped to the surface. The most expensive raw material is ammonia, which is made by the reaction of elemental hydrogen and nitrogen over an iron-based catalyst,

$$3\,H_2 + N_2 \rightarrow 2\,NH_3$$

a means of making ammonia developed by Haber and Bosch in Germany in 1913. However, as shown below, the ammonia is recycled, so only relatively small quantities of additional makeup NH_3 are required.

In addition to NaCl, the major consumable raw material in the Solvay process is calcium carbonate, $CaCO_3$, which is abundantly available from deposits of limestone. It is heated (calcined)

$$CaCO_3 + \text{heat} \rightarrow CaO + CO_2$$

to produce calcium oxide and carbon dioxide gas. The carbon dioxide gas is used in Reaction 5.13.1, another green chemical aspect of the process. The calcium oxide is reacted with water (it is said to be slaked),

$$CaO + H_2O \rightarrow Ca(OH)_2$$

to produce basic calcium hydroxide. This base is then reacted with the solution from which solid $NaHCO_3$ has been precipitated (Reaction 5.13.1) and that contains dissolved ammonium chloride,

$$Ca(OH)_2\,(s) + 2\,NH_4Cl\,(aq) \rightarrow 2\,NH_3\,(g) + CaCl_2\,(aq) + 2\,H_2O\,(l)$$

releasing ammonia gas that is recycled back into Reaction 5.13.1 for $NaHCO_3$ synthesis. This has the advantage of recycling ammonia, which is essential for the process to be economical. It has the disadvantage of generating a solution of calcium chloride, $CaCl_2$. The commercial demand for this salt is limited, although concentrated solutions of it are used for de-icing ice-covered roads. It has such a voracious appetite for water that it cannot be dried economically for storage in a dry form.

Does the Solvay process meet the criteria for a green chemical synthesis? There is not a simple answer to that question. There are two respects in which it does meet green chemical criteria:

1. It uses inexpensive, abundantly available raw materials in the form of NaCl brine and limestone ($CaCO_3$). A significant amount of NH_3 is required to initiate the process with relatively small quantities to keep it going.

2. It maximizes recycle of two major reactants, ammonia and carbon dioxide. The calcination of limestone (Reaction 5.13.4) provides ample carbon dioxide to make up for inevitable losses from the process, but some additional ammonia has to be added to

compensate for any leakage.

What about the percent yield and atom economy of the Solvay process? The percent yield of reaction generating the product, Reaction 5.13.1, can be expected to be significantly less than 100% in large part because the stoichiometric amount of $NaHCO_3$ cannot be expected to precipitate from the reaction mixture. To calculate the maximum atom economy for Na_2CO_3 production, it must be assumed that all reactions go to completion without any losses. In such an ideal case, the overall reaction for the process is

$$CaCO_3 + 2\,NaCl \rightarrow Na_2CO_3 + CaCl_2$$

Using the atomic masses Ca 40.0, C 12.0, O 16.0, and Cl 35.5 gives the molar masses of $CaCO_3$, 100 g/mol; NaCl, 58.5 g/mol; Na_2CO_3, 106 g/mol; and $CaCl_2$, 111 g/mol. If the minimum whole number of moles of reactants were to react, 100 g of $CaCO_3$ would react with 2×58.5 = 117 g of NaCl to produce 106 g of Na_2CO_3 and 111 g of $CaCl_2$. Note that the mass of NaCl reacting is 2 times the molar mass because 2 moles of NaCl are reacting. So, for these amounts of materials inthe reaction, a total mass of 100 + 117 = 217 g of reactants produces 106 g of the Na_2CO_3 product. Therefore, the percent atom economy is

$$\text{Percent atom economy} = \frac{\text{Mass of desired product}}{\text{Total mass of reactants}} \times 100 = \frac{106g}{217g} \times 100 = 48.8\%$$

This is the maximum possible value assuming complete reactions and no losses. If the CaCl2byproduct is considered to be a useful product, the atom economy can be regarded as being higher.

Is the Solvay process green with respect to environmental impact? Again, the answer to this question is mixed. Extraction of the two major raw materials, limestone and NaCl, normally can be accomplished with minimal adverse effects on the environment. Quarrying of limestone in open pits results in dust production and blasting of the rock, which is usually carried out with an explosive mixture of fuel oil mixed with ammonium nitrate, NH_4NO_3, causes some disturbance. Open-pit lime stone quarries can be unsightly, but can also serve as artificial lakes. In some places the underground spaces left from the underground quarrying of limestone have found excellent commercial use as low-cost warehouses that largely provide their own climate control. Truck transport of quarried lime definitely has negative environmental impacts. Extraction of liquid NaCl brine usually has minimal environmental impact. The Solvay process, itself, releases significant quantities of greenhouse gas CO_2 and some gaseous ammonia to the atmosphere. Solvay production of sodium carbonate requires significant amounts of energy.

There are numerous natural deposits of sodium bicarbonate and sodium carbonate. The most common source of these salts is a mineral called **trona**, for which the chemical formula is Na_2CO_3•$NaHCO_3$•$2H_2O$. (This formula shows that a formula unit of trona mineral consists of 1 formula unit of ionic Na_2CO_3, 1 formula unit of ionic $NaHCO_3$ and 2 molecules of H_2O). The development of huge deposits of trona in the state of Wyoming and elsewhere in the world has lowered dependence on the Solvay process for sources of sodium bicarbonate and sodium carbonate and the process is no longer used in the United States.

This page titled 2.5.5.13: Industrial Chemical Reactions - The Solvay Process is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **5.13: Industrial Chemical Reactions - The Solvay Process** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.5.14: Questions and Problems

1. How do chemical equations relate to chemical reactions?

2. Summarize the information contained in the chemical equation below. How would this reaction be classified? $CaCl_2(aq) + Na_2CO_3(aq) \rightarrow CaCO_3(s) + 2NaCl(aq)$

3. What are the meanings of (*s*), (*l*), (*g*), and (*aq*) after formulas in a chemical equation? What are the meanings of Δ and $\leftarrow \rightarrow$?

4. What is wrong with balancing the chemical equation $S + O_2 \rightarrow SO_3$ as $S + O_2 \rightarrow SO_2$?

5. From your knowledge of chemistry and chemical formulas write the balanced equation for heating magnesium carbonate to give magnesium oxide and carbon dioxide, indicating the physical states of the reactants and products.

6. Balance the equation $FeSO_4 + H_2SO_4 + O_2 \rightarrow Fe_2(SO_4)_3 + H_2O$, which is for a reaction involved in the formation of pollutant acid mine water.

7. Balance each of the following: (a) $C_2H_4 + O_2 \rightarrow CO_2 + H_2O$, (b) $KClO_4 \rightarrow KClO + O_2$, (c)$FeS_2 + O_2 + H_2O \rightarrow FeSO_4 + H_2SO_4$ (d) $Fe_2O_3 + CO \rightarrow Fe + CO_2$, (e) $H_3PO_4 + H2 \rightarrow PH_3 + H_2O$, (f) $P + Cl_2 \rightarrow PCl_5$

8. Explain how chemical equations fit in with the general scheme of chemistry as a language.

9. A chemical equation that describes the action of hydrogen sulfide, H2S, dissolved in water is$H2S \leftarrow \rightarrow H+ + HS-$. What does this equation say and how is it consistent with the fact that dissolved hydrogen sulfide is a weak acid?.

10. From the discussion of reactions of metals with sulfuric acid in Section 5.3 and your knowledge of the properties of silver jewelry, explain what is likely to happen when silver metal is placed in sulfuric acid.

11. Zinc is a very reactive metal. Explain with chemical equations what you would expect to happen if zinc metal were placed in sulfuric acid and what would happen if zinc oxide, ZnO, were placed in sulfuric acid.

12. Finely divided steel wool heated red hot and quickly placed into a bottle of oxygen burns vigorously undergoing the reaction $4Fe + 3O_2 \rightarrow 2Fe_2O_3$. Why is there no concern that a steel beam used in construction will burn in air? However, such a beam can be cut with an oxyacetylene torch by first heating a small portion of it with the torch, then turning off the acetylene and slowly running the torch across the beam. What is happening in this case?

13. A water solution of hydrogen peroxide, H_2O_2, is relatively stable. But, if a small quantity of solid manganese oxide is placed in the solution of hydrogen peroxide, bubbles are given off near the surface of the manganese oxide, although the solid appears to remain intact. Explain what happens and the role of the manganese oxide.

14. The following reactions were given in connection with the Solvay process used to makesodium bicarbonate and sodium carbonate: (A) $NaCl + NH3 + CO_2 + H_2O \rightarrow NaHCO_3(s) + NH_4Cl$, (B) $2NaHCO_3 + heat \rightarrow Na_2CO_3 + H_2O(g) + CO2(g)$ (C) $3H_2 + N_2 \rightarrow 2NH_3$, (D) $CaCO_3 + heat \rightarrow CaO + CO_2$, (E) $CaO + H_2O \rightarrow Ca(OH)_2$. Classify each of these reactions in the categories given in Section 5.6.

15. Given the chemical reaction $4CH_4 + 6NO_2 \rightarrow 4CO + 3N_2 + 8H_2O$, write all the possible mole ratios relating N_2 to each of the other reaction participants.

16. Given the atomic masses N 14.0, H 1.0, and Cl 35.5 and the reaction below, calculate the mass of HCl produced when 12.7 g of NH_3 react.

$2NH_3 + 3Cl_2 \rightarrow N_2 + 6HCl$

17. Given the atomic masses C 12.0, H 1.0, and O 16.0 and the reaction below, calculate the mass of H_2O produced when 15.6 g of O_2react.

$C_2H_4 + 3O_2 \rightarrow 2CO_2 + 2H_2O$

18. Match the reaction type from the list on the left with the example reaction from the right, below. $PbSO_4$ is insoluble in water.

A. Decomposition 1. $HCl + NaOH \rightarrow H2O + NaCl$

B. Neutralization 2. $Pb(NO_3)_2 + Na_2SO_4 \rightarrow PbSO_4 + 2NaCl$

C. Substitution 3. $2H_2O_2 \rightarrow 2H_2O + O_2$

D. Double displacement 4. $CuSO_4(aq) + Fe(s) \rightarrow FeSO_4(aq) + Cu(s)$

19. Of the following, the **untrue** statement is

A. The symbol $\leftarrow \rightarrow$ is used to show that a reaction goes both ways.

B. The notation (l) is used to show that a reactant or product is dissolved in water.

C. A catalyst changes the rate of a reaction but is not itself consumed.

D. The symbol, Δ, is used to show application of heat to a reaction.

E. Simply because a chemical equation may be written and balanced does **not** indicate for certain that the chemical reaction it indicates will occur.

20. Given the reaction $2S + 3O_2 \rightarrow 2SO_2$ and atomic masses of 32.0 and 16.0 for S and O, respectively, calculate the mass of O_2 reacting with 15.0 g of S.

21. Given the reaction $CH_4 + 2H_2O \rightarrow CO_2 + 4H_2$ and atomic masses of C, 12.0, H, 1.0; and O16.0, calculate the total mass of products formed when 24.0 g of CH_4 reacts.

22. Given the reaction $3CH_4 + 4Fe_2O_3 \rightarrow 3CO_2 + 6H_2O + 8Fe$ and atomic masses of C, 12.0; H,1.0; Fe, 55.8; and O, 16.0, what is the mass of CO_2 produced by the reaction of 36.0 g of Fe_2O_3?

23. What is the basis of stoichiometry in respect to relative amounts of materials in reactions?

23. What are the major steps in doing a stoichiometric calculation?

24. What is a limiting reactant?

25. A solution of $FeSO_4$ was prepared by mixing 100 g of pure H_2SO_4 with water and putting it in contact with 50.0 g of iron metal. What reaction occurred? What masses of reaction products were generated and what were the masses of reactants, if any, left over? The atomic masses needed are H 1.0, Fe 55.8, S 32.0, and O 16.0.

26. What is the difference between the stoichiometric yield and the measured yield in a chemical reaction? How are they used to calculated percent yield?

27. How are titrations and stoichometry related?

28. A solid mineral sample consisting of calcium carbonate, $CaCO_3$, and nonreactive mineral matter weighing 0.485 g was stirred in some water to which 0.115 mol/L standard hydrochloric acid, HCl, was added from a buret. The reaction was $CaCO_3 + 2HCl \rightarrow CaCl_2 + CO_2 + H_2O$. If 48.6 milliliters (0.0486 L) of HCl was required to react with all the $CaCO_3$ in the sample, what was the percentage of $CaCO_3$ in the sample given that the molar mass of $CaCO_3$ is 100 g/mol?

29. A 250 mL sample of incinerator exhaust gas scrubber water contaminated with HCl was titrated with 0.104 mol/L standard NaOH, of which 11.3 mL were required to reach the endpoint. What was the molar concentration of HCl in the water sample?

30. What is made by the Solvay process? What is the overall chemical reaction that describes the Solvay process? What are the two major raw materials consumed and what are two major species that are recycled through the process?

31. What are major green aspects of the Solvay process? What are some aspects that are less green?

32. What is a major alternative to use of the Solvay process?

33. Calculate the number of moles of total of $AlCl_3$ in 38.6 g of the compound and the number of moles of CH_4, in 217 g of methane. Use 27.0, 35.5, 12.0, and 1.0 for the atomic masses of Al, Cl, C, and H, respectively.

34. Why might you expect stoichiometric ratios of reactants to be used in industrial chemical reactions? If one of two reactants used in an industrial process is much more expensive than another, suggest why and in which way a stoichiometric ratio might not be used? Also, suppose that one of two reactants is quite toxic whereas the other reactant is not. Why might the practice of green chemistry suggest using a nonstoichometric ratio of reactants in such a case?

35. Given the reaction $2H_2 + O_2 \rightarrow 2H_2O$, identify which species is oxidized, which is reduced, which is the oxidizing agent, and which is the reducing agent.

36. Given the reaction that occurs when a direct electrical current is passed through liquid ionic NaCl, $2Na^+ + 2Cl^- \rightarrow 2Na + Cl_2$, identify which species is oxidized and which is reduced. Justify the answer.

37. Identify which reactions given in Section 5.6 are oxidation-reduction reactions

---

This page titled 2.5.5.14: Questions and Problems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **Questions and Problems** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

2.5.6: The Wonderful World of Carbon - Organic Chemistry and Biochemicals

“The first few of what are now known to be organic chemicals to be discovered were produced by living organisms and were therefore called ‘organic.’ One such compound is urea, which occurs in urine. In 1828 Friedrich Wöhler disproved the idea that all organic compounds must come from living organisms when he accidentally discovered that urea could be made by the reaction of cyanic acid (HOCN) and ammonia, both simple organic compounds. This discovery established the science of organic chemistry based upon the unique bonding capabilities of the carbon atom leading to the synthesis and discovery of tens of millions of unique organic compounds. In 2009 the American Chemical Society Chemical Abstract Service registered the 60 millionth compound(most of which are organic compounds) only 9 months after the registration of the 40 millionth, apace of discovery of more than one new compound per minute.”

2.5.6.1: Rings and Chains of Carbon Atoms

2.5.6.2: Compounds of Carbon and Hydrogen - Hydrocarbons

2.5.6.3: Using Lines To Show Organic Structural Formulas

2.5.6.4: Functional Groups

2.5.6.5: Giant Molecules from Small Organic Molecules

2.5.6.6: Questions and Problems

2.5.6.7: Supplementary References

This page titled 2.5.6: The Wonderful World of Carbon - Organic Chemistry and Biochemicals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.6.1: Rings and Chains of Carbon Atoms

Most of the molecules of chemical compounds studied so far have been clusters of only a few atoms. Therefore, molecules of water, H_2O, exist as individual clusters of 2 H atoms bonded to 1O atom and molecules of ammonia, NH_3, each consist of an atom of N to which are bonded 3 H atoms. In cases where atoms of a particular element in chemical compounds have a tendency to bond with atoms of the same element, the number of possible compounds is increased tremendously. This is the case with carbon, C. Groups of carbon atoms can bond together to form straight chains, branched chains, and rings, leading to a virtually limitless number of chemical compounds. Such carbon-containing compounds are **organic chemicals**, the study of which is **organic chemistry**. Adding to the enormous diversity of organic chemistry is the fact that two carbon atoms may be connected by **singlebonds** consisting of 2 shared electrons, **double bonds** composed of 4 shared electrons, and even **triplebonds** that contain 6 shared electrons.

Organic chemicals comprise most of the substances with which chemists are involved. Petroleum, which serves as the raw material for vast polymer, plastics, rubber, and other industries consists of hundreds of compounds composed of hydrogen and carbon called **hydrocarbons**.Among organic chemicals are included the majority of important industrial compounds, synthetic polymers, agricultural chemicals, and most substances that are of concern because of their toxicities and other hazards. The carbohydrates, proteins, lipids (fats and oils), and nucleic acids(DNA) that make up the biomass of living organisms are organic chemicals made by biological processes. The feedstock chemicals needed to manufacture a wide range of chemical products are mostly organic chemicals, and their acquisition and processing are of great concern in the practice of green chemistry. The largest fraction of organic chemicals acquired from petroleum and natural gas sources are burned to fuel vehicles, airplanes, home furnaces, and power plants. Prior to burning, these substances may be processed to give them desired properties. This is particularly true of the constituents of gasoline, the molecules of which are processed and modified to give gasoline desired properties of smooth burning (good antiknock properties) and low air pollution potential. Pollution of the water, air, and soil environments by organic chemicals is an area of significant concern. Much of the effort put into green chemistry has involved the safe manufacture, recycling, and disposal of organic compounds.

A number of organic compounds are made by very sophisticated techniques to possess precisely tailored properties. This is especially true of pharmaceuticals, which must be customized to deliver the desired effects with minimum undesirable side effects. A single organic compound that is effective against one of the major health problems — usually one out of hundreds or even thousands tested — has the potential for hundreds of millions of dollars per year in profits.

Organic chemicals differ widely in their toxicities. Some compounds are made and used because of their toxicities to undesirable organisms. These are the **pesticides**, including, especially, **insecticides** used to kill unwanted insects and **herbicides** used to eradicate weeds that compete with desired crops. Green chemistry is very much involved with these kinds of applications. One of the more widely applied uses of genetically modified crops has been the development of crops that produce their own insecticides in the form of insecticidal proteins normally made by certain kinds of bacteria whose genes have been spliced into field crops. Another application of green chemistry through genetic engineering is the development of crops that resist the effects of specific organic molecules commonly used as herbicides. These herbicides may be applied directly to target crops, leaving them unscathed while competing weeds are killed.

It should be obvious from this brief discussion that organic chemistry is a vast, diverse, highly useful discipline based upon the unique bonding properties of the carbon atom. The remainder of this chapter discusses major aspects of organic chemistry. Many of the most interesting and important organic chemicals are made by biological processes. Indeed, until 1828, it was generally believed that only organisms could synthesize organic chemicals. In that year, Friedrich Wöhlers succeeded in making urea, an organic chemical that is found in urine, from ammonium cyanate, an inorganic material. Because of the important role of organisms in making organic chemicals, several of the most significant kinds of these chemicals made biologically are also discussed in this chapter. Additional details regarding the ways in which living organisms make and process chemicals are given in Chapters 9 and 13.

This page titled 2.5.6.1: Rings and Chains of Carbon Atoms is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **6.1: Rings and Chains of Carbon Atoms** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.6.2: Compounds of Carbon and Hydrogen - Hydrocarbons

The tremendous variety and diversity of organic chemistry is due to the ability of carbon atoms to bond with each other in a variety of straight chains, branched chains, and rings and of adjacent carbon atoms to be joined by single, double, or triple bonds. This bonding ability can be illustrated with the simplest class of organic chemicals, the **hydrocarbons** consisting only of hydrogen and carbon. Figure 6.1 shows some hydrocarbons in various configurations. Hydrocarbons are the major ingredients of petroleum and are pumped from the ground as crude oil or extracted as natural gas. They have two major uses. The first of these is combustion as a source of fuel. The most abundant hydrocarbon in natural gas, methane, CH_4, is burned in home furnaces, electrical power plants, and even in vehicle engines,

$$CH_4 + 2\,O_2 \rightarrow CO_2 + 2\,H_2O + \text{heat energy}$$

to provide energy. The second major use of hydrocarbons is as a raw material for making rubber, plastics, polymers, and many other kinds of materials. Given the value of hydrocarbons as a material, it is unfortunate that so much of hydrocarbon production is simply burned to provide energy, which could be generated by other means.

There are several major class of hydrocarbons, all consisting of only hydrogen and carbon. **Alkanes** have only single bonds between carbon atoms. Cyclohexane, *n*-heptane, and 3-ethyl-2,5-dimethylhexane in Figure 6.1 are alkanes; the cyclohexane is a cyclic hydrocarbon. **Alkenes**, such as propene shown in Figure 6.1, have at least one double bond consisting of 4 shared electrons between two of the carbon atoms in the molecule. **Alkynes** have at least one triple bond between carbon atoms in the molecule as shown for acetylene in Figure 6.1. Acetylene is an important fuel for welding and cutting torches; otherwise, the alkynes are of relatively little importance and will not be addressed farther. A fourth class of hydrocarbon consists of **aromatic** compounds which have rings of carbon atoms with special bonding properties as discussed later in this chapter

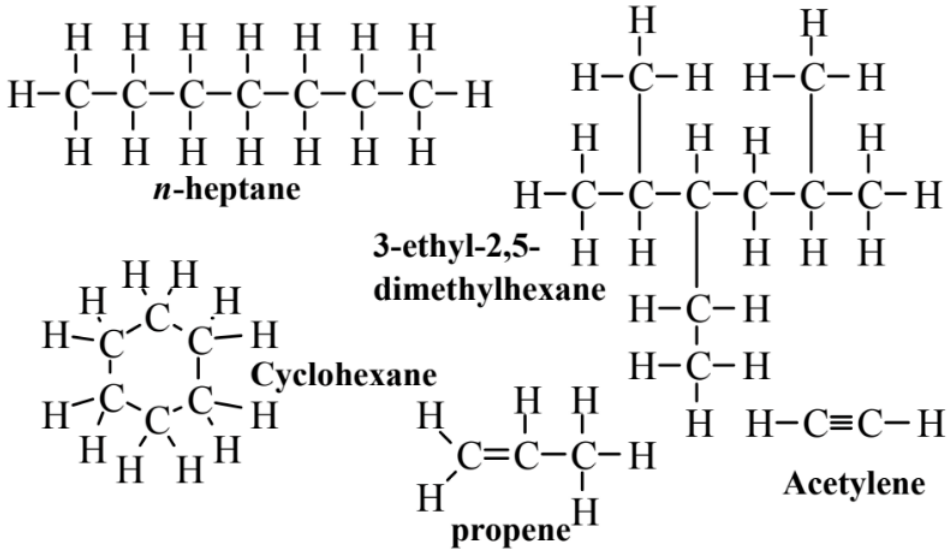


Figure 6.1. Some typical hydrocarbons. These formulas illustrate the bonding diversity of carbon which gives rise to an enormous variety of hydrocarbons and other organic compounds.

Alkanes

The molecular formulas of non-cyclic alkanes are CnH2n+2. By counting the numbers of carbon and hydrogen atoms in the molecules of alkanes shown in Figure 6.1, it is seen that the molecular formula of *n*-heptane is C_7H_6 and that of 3-ethyl-2,5-dimethylhexane is $C_{10}H_{22}$, both of which fit the general formula given above. The general formula of cyclic alkanes is C_nH_{2n}; that of cyclohexane, the most common cyclic alkane, is C_6H_{12}. These formulas are **molecular formulas**, which give the number of carbon and hydrogen atoms in each molecule, but do not tell anything about the structure of the molecule. The formulas given in Figure 6.1 are **structuralformulas** which show how the molecule is assembled. The structure of *n*-heptane is that of a straight chain of carbon atoms; each carbon atom in the middle of the chain is bound to 2 H atoms and the 2 carbon atoms at the ends of the chain are each bound to 3 H atoms. The prefix *hep* in the name denotes 7 carbon atoms and then*-*indicates that the compound

consists of a single straight chain. This compound can be represented by a **condensed structural formula** as $CH_3(CH_2)_5CH_3$ representing 7 carbon atoms in a straight chain. In addition to methane mentioned previously, the lower alkanes include the following:

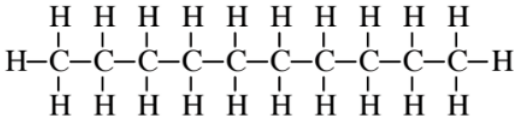
Ethane: CH_3CH_3 Propane: $CH_3CH_2CH_3$ Butane: $CH_3(CH_2)_2CH_3$ *n*-Pentane: $CH_3(CH_2)3CH_3$

For alkanes with 5 or more carbon atoms, the prefix (*pen* for 5,*hex* for 6,*hept* for 7,*oct* for 8,*non* for 9) shows the total number of carbon atoms in the compound and *n*-may be used to denote a straight-chain alkane. Condensed structural formulas may be used to represent branched chain alkanes as well. The condensed structural formula of 3-ethyl-2,5-dimethylhexane is

$$CH_3CH(CH_3)CH(C_2H_5)CH_2CH(CH_3)CH_3$$

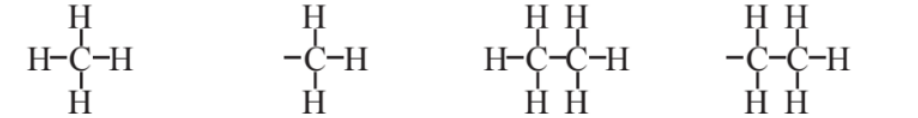
In this formula, the C atoms and their attached H atoms that are not in parentheses show carbons that are part of the main hydrocarbon chain. The (CH_3) after the second C in the chain shows a methyl group attached to it, the (C_2H_5) after the third carbon atom in the chain shows an ethyl group attached to it, and the (CH_3) after the fifth carbon atom in the chain shows a methyl group attached to it.

Compounds that have the same molecular formulas but different structural formulas are **structural isomers**. For example, the straight-chain alkane with the molecular formula $C_{10}H_{22}$ is *n*-decane



which is a structural isomer of 3-ethyl-2,5-dimethylhexane.

The names of organic compounds are commonly based upon the structure of the hydrocarbon from which they are derived using the longest continuous chain of carbon atoms in the compound as the basis for the name. For example, the longest continuous chain of carbon atoms in 3-ethyl-2,5-dimethylhexane shown in Figure 6.1 is 6 carbon atoms, so the name is based upon *hex*ane. The names of the chain branches are also based upon the alkanes from which they are derived. As shown below,



the two shortest-chain alkanes are methane with 1 carbon atom and ethane with 2 carbon atoms. Removal of 1 of the H atoms from methane gives the **methyl** group and removal of 1 of the H atoms from ethane gives the **ethyl** group. These terms are used in the name 3-ethyl-2,5-dimethylhexane to show groups attached to the basic hexane chain. The carbon atoms in this chain are numbered sequentially from left to right. An ethyl group is attached to the 3rd carbon atom, yielding the “3-ethyl” part of the name, and methyl groups are attached to the 2nd and 5th carbon atoms, which gives the “2,5-dimethyl” part of the name.

The names discussed above are **systematic names**, which are based upon the actual structural formulas of the molecules. In addition, there are **commonnames** of organic compounds that do not indicate the structural formulas. Naming organic compounds is a complex topic, and no attempt is made here to teach it to the reader. However, from the names of compounds given in this and later chapters, some appreciation of the rationale for organic compound names should be obtained.

Other than burning them for energy, the major kind of reaction with alkanes consists of **substitution reactions** such as,

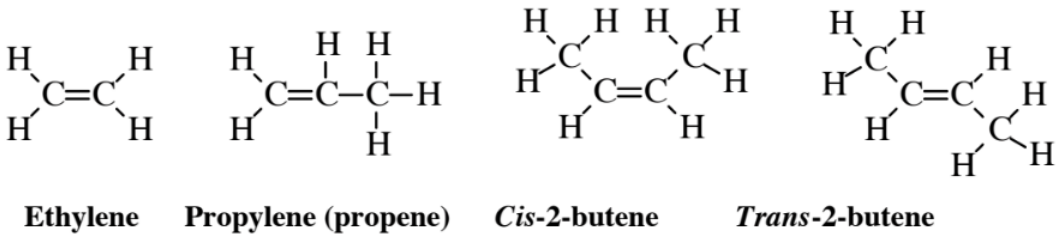
$$C_2H_6 + 2\,Cl_2 \rightarrow C_2H_4Cl_2 + 2\,HCl$$

in which one or more H atoms are displaced by another kind of atom. This is normally the first step in converting alkanes to compounds containing elements other than carbon or hydrogen for use in synthesizing a wide variety of organic compounds.

Alkenes

Four common alkenes are shown in Figure 6.2. Alkenes have at least one C=C double bond per molecule and may have more. The first of the alkenes in Figure 6.2, ethylene, is a very widely produced hydrocarbon used to synthesize polyethylene plastic and other

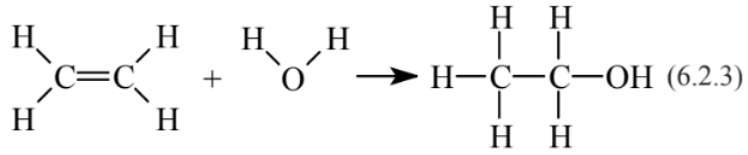
organic compounds. About 25 billion kilograms (kg) of ethylene are processed in the U.S. each year. About 14.5 billion kg of propylene are used in the U.S. each year to produce polypropylene plastic and other chemicals. The two 2-butene compounds illustrate an important aspect of alkenes, the possibility of ***cis*-*trans***isomerism. Whereas carbon atoms and the groups substituted onto them joined by single bonds can freely rotate relative to each other as though they were joined by a single shaft, carbon atoms connected by a double bond behave as though they were attached by two parallel shafts and are not free to rotate. So, *cis*-2-butene in which the two end methyl (-CH3) groups are on the same side of the molecule is a different compound from *trans*-2-butene in which they are on opposite sides. These two compounds are *cis*-*trans* isomers



Ethylene **Propylene (propene)** ***Cis*-2-butene** ***Trans*-2-butene**

Figure 6.2. Examples of alkene hydrocarbons

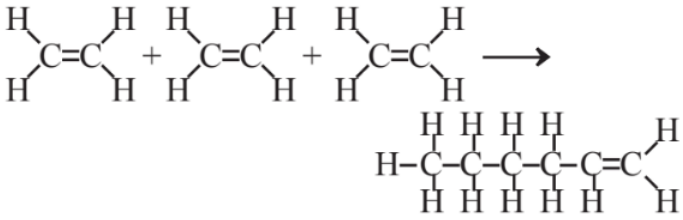
Alkenes are chemically much more active than alkanes. This is because the double bond is **unsaturated** and has electrons available to form additional bonds with other atoms. This leads to **addition reactions** in which a molecule is added across a double bond. For example, the addition of H_2O to ethylene



Ethylene **Water** **Ethanol**

yields ethanol, the same kind of alcohol that is in alcoholic beverages. In addition to adding immensely to the chemical versatility of alkenes, addition reactions make them quite reactive in the atmosphere during the formation of photochemical smog. The presence of double bonds also adds to the biochemical and toxicological activity of compounds in organisms

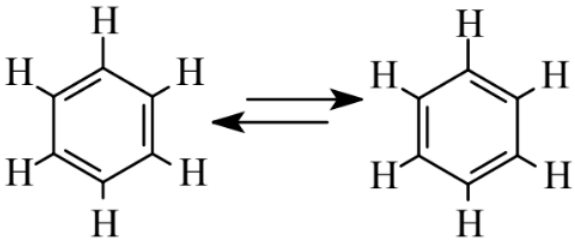
Because of their double bonds, alkenes can undergo **polymerization** reactions in which large numbers of individual molecules add to each other to produce big molecules called **polymers** (see Section 6.5). For example, 3 ethylene molecules can add together as follows:



a process that can continue, forming longer and longer chains and resulting in the formation of huge molecules molecules of polyethylene

Aromatic Hydrocarbons

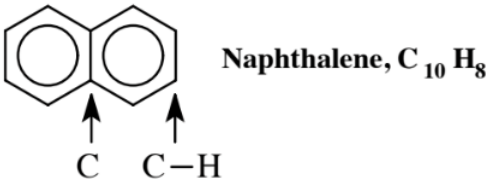
A special class of hydrocarbons consists of rings of carbon atoms, almost always containing 6C atoms, which can be viewed as having alternating single and double bonds as shown below:



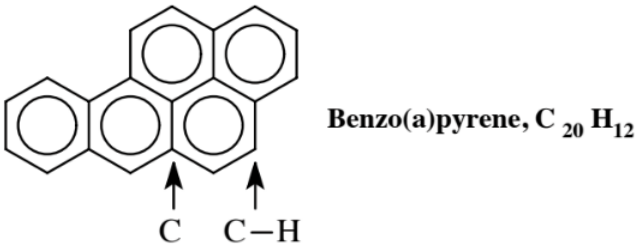
These structures show the simplest aromatic hydrocarbon, benzene, C6H6. Although the benzene molecule is represented with 3 double bonds, chemically it differs greatly from alkenes, for example undergoing substitution reactions rather than addition reactions. The properties of aromatic compounds are special properties called **aromaticity**. The two structures shown above are equivalent **resonance** structures, which can be viewed as having atoms that stay in the same places, but in which the bonds joining the atoms can shift positions with the movement of electrons composing the bonds. Since benzene has different chemical properties from those implied by either of the above structures, it is commonly represented as a hexagon with a circle in the middle:



Many aromatic hydrocarbons have two or more rings. The simplest of these is naphthalene,



a two-ringed compound in which two benzene rings share the carbon atoms at which they are joined; these two carbon atoms do not have any H attached, each of the other 8 C atoms in the compound has 1 H attached. Aromatic hydrocarbons with multiple rings, called **polycyclic aromatic hydrocarbons**, PAH, are common and are often produced as byproducts of combustion. One of the most studied of these is benzo(a)pyrene,



found in tobacco smoke, diesel exhaust, and charbroiled meat. This compound is toxicologically significant because it is partially oxidized by enzymes in the body to produce a cancer-causing metabolite.

The presence of hydrocarbon groups and of elements other than carbon and hydrogen bonded to an aromatic hydrocarbon ring gives a variety of **aromatic compounds**. Three examples of common aromatic compounds are given below. Toluene is widely used for chemical synthesis and as a solvent. The practice of green chemistry now calls for substituting toluene for benzene wherever possible because benzene is suspected of causing leukemia, whereas the body is capableof metabolizing toluene to harmless metabolites (see Chapter 7). About 850 million kg of aniline are made in the U.S. each year as an intermediate in the synthesis of dyes and other organic chemicals. Phenol is a relatively toxic oxygen-containing aromatic compound which, despite its toxicity to humans, was the first antiseptic used in the 1800s.



This page titled 2.5.6.2: Compounds of Carbon and Hydrogen - Hydrocarbons is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **6.2: Compounds of Carbon and Hydrogen - Hydrocarbons** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.6.3: Using Lines To Show Organic Structural Formulas

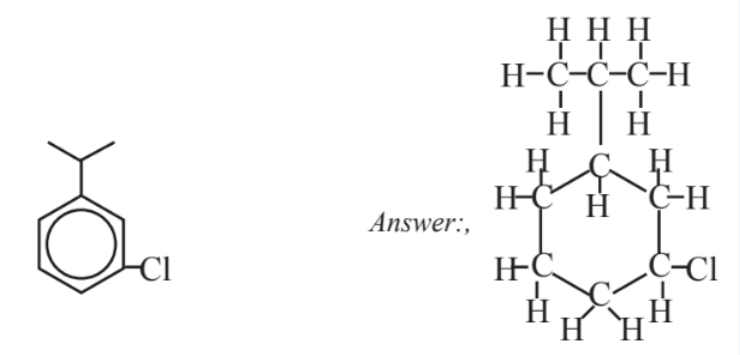
The aromatic structures shown above use a hexagon with a circle in it to denote an aromatic benzene ring. Organic chemistry uses lines to show other kinds of structural formulas as well. The reader who may have occasion to look up organic formulas will probably run into this kind of notation, so it is important to be able to interpret these kinds of formulas. Some line formulas are shown in Figure 6.3.

In using lines to represent organic structural formulas, the corners where lines intersect and the ends of lines represent C atoms, and each line stands for a covalent bond (2 shared electrons). It is understood that each C atom at the end of a single line has 3 H atoms attached, each C atom at the intersection of 2 lines has 2 C atoms attached, each C at the intersection of 3 lines has 1 H attached, and the intersection of 4 lines denotes a C atom with no H atoms attached. Multiple lines represent multiple bonds as shown for the double bonds in 1,3-butadiene. Substituent groups are shown by their symbols (for individual atoms), or formulas of functional groups consisting of groups of atoms; it is understood that each such group substitutes for a hydrogen atom as shown in the formula of 2,3-dichlorobutane in Figure 6.3. The 6-carbon-atom aromatic ring is denoted by a hexagon with a circle in it.

Exercise

What is the structural formula of the compound represented on the left, below?

Answer



This page titled 2.5.6.3: Using Lines To Show Organic Structural Formulas is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **6.3: Using Lines To Show Organic Structural Formulas** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.6.4: Functional Groups

Numerous elements in addition to carbon and hydrogen occur in organic compounds. These are contained in **functional groups**, which define various classes of organic compounds. The $-NH_2$ group in aniline and the -OH groups in phenol mentioned above are examples of functional groups. The same organic compound may contain two or more functional groups. Among the elements common in functional groups are O, N, Cl, S, and P. There is not space here to discuss all the possible functional groups and the classes of organic compounds that they define. Some important examples are given to provide an idea of the variety of organic compounds with various functional groups. Other examples are encountered later in the text

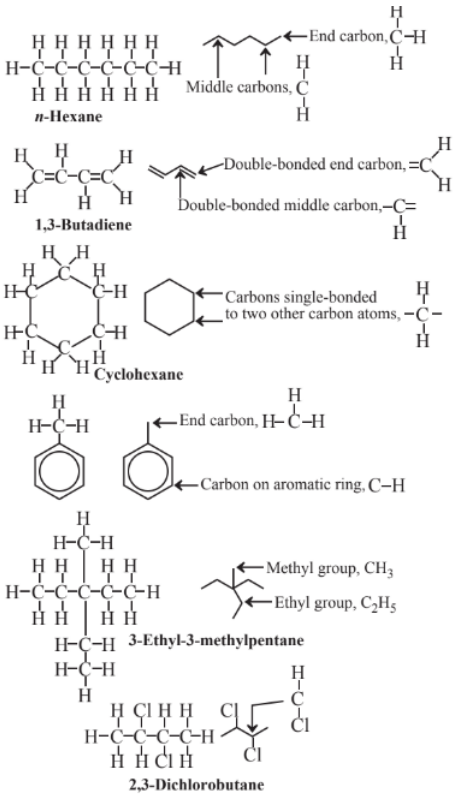
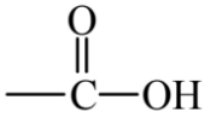


Figure 6.3. Representation of organic structural formulas with lines. The structural formulas showing the atoms are on the left and the corresponding line formulas on the right. Each line corner and end represents a carbon atom unless otherwise specified. Each C atom has four covalent bonds or equivalent attached and the presence of H atoms is implied to provide the required bonds.

In using lines to represent organic structural formulas, the corners where lines intersect and the ends of lines represent C atoms, and each line stands for a covalent bond (2 shared electrons). It is understood that each C atom at the end of a single line has 3 H atoms attached, each C atom a the intersection of 2 lines has 2 C atoms attached, each C at the intersection of 3 lines has 1 H attached, and the intersection of 4 lines denotes a C atom with no H atoms attached. Multiple lines

Organooxygen Compounds

Figure 6.4 shows several important classes of organic compounds that contain oxygen. **Ethylene oxide** is a sweet-smelling, colorless, flammable, explosive gas. It is an **epoxide** characterized by an oxygen atom bridging two carbon atoms that are also bonded with each other. Ethylene oxide is toxic and is used as a sterilant and fumigant as well as a chemical intermediate. Because of the toxicity and flammability of this compound, the practice of green chemistry tries to avoid its generation and use. **Ethanol**, which occurs in alcoholic beverages, is an **alcohol**, a class of compound in which the -OH group is bonded to an alkane or alkene (attachment of the -OH group to an aromatic hydrocarbon molecule gives a **phenolic** compound). **Acetone** is a ketone, a class of compounds that has the C=O functional group in the middle of a hydrocarbon chain. Acetone is an excellent organic solvent and relatively safe. **Butyric** acid, which occurs in butter, is an organic carboxylic acid, all of which contain the functional group,



which can release the H+ion characteristic of acids. **Methyltertiarybutyl ether**, MTBE, is an example of an ether in which an O atom connects 2 C atoms. When highly toxic tetraethyllead was phased out of gasoline as an octane booster, MTBE was chosen as a substitute. It was subsequently found to be a particularly noxious water pollutant, and its use has been largely banned

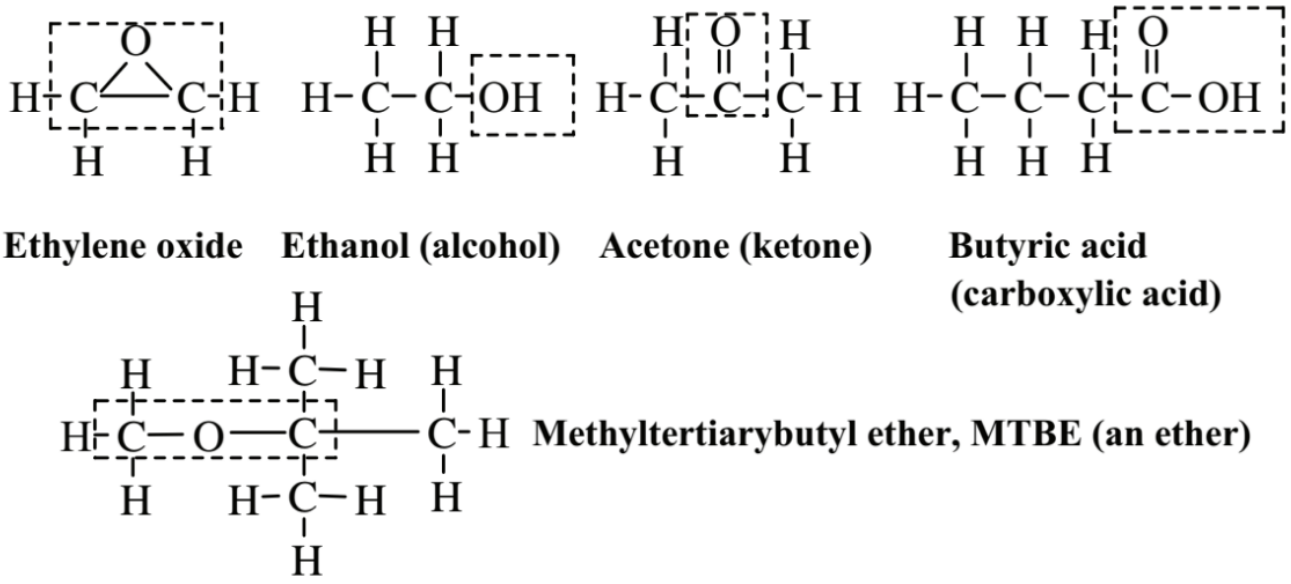
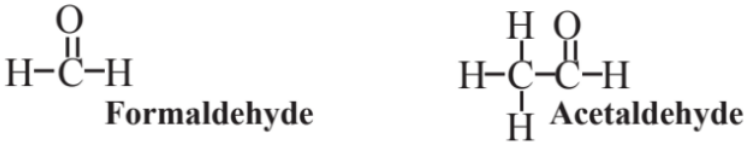


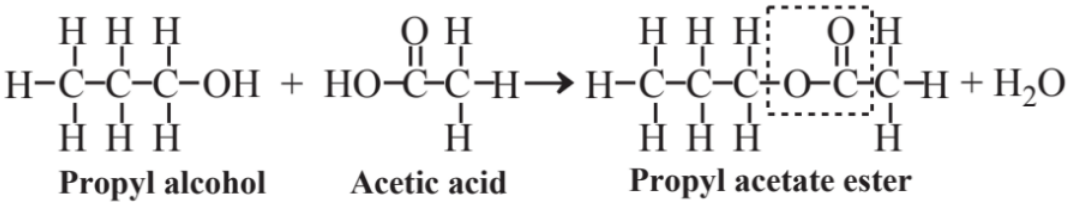
Figure 6.4. Examples of important classes of organic compound with oxygen-containing functional groups. The functional groups characteristic of various classes of compounds are outlined by dashed lines.

The C=O group in the middle of an organic molecule is characteristic of ketones. When this group is located at the end of a molecule and the carbon is also bonded to H, the compound is an **aldehyde**. The two lowest aldehydes are formaldehyde and acetaldehyde,



of which formaldehyde is the most widely produced. Despite its many uses, formaldehyde lacks characteristics of green chemicals because it is a volatile, toxic, noxious substance. Formaldehyde tends to induce hypersensitivity (allergies) in people who inhale the vapor or whose skin is exposed to it.

The reaction of an alcohol and an organic acid,



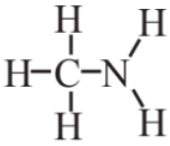
(6.4.1)

produces an important kind of organic compound called **esters**. The linkage characteristic of esters is outlined by the dashed box in the structure of propyl acetate above. A large number of the naturally-occurring esters made by plants are noted for their pleasant

odors. Propyl acetate, for example gives pears their pleasant odor. Other fruit odors due to esters include methyl butyrate, apple; ethyl butyrate, pineapple; and methyl benzoate, ripe kiwi fruit.

Organonitrogen Compounds

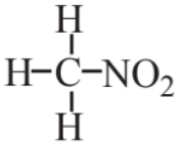
Methylamine,



Methylamine

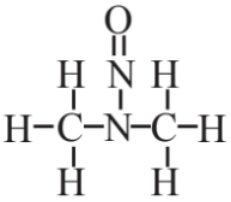
is the simplest of the amines, compounds in which an N atom is bonded to a hydrocarbon group. In an amine, the N atom may be bonded to 2 H atoms, or one or both of these H atoms may be substituted by hydrocarbon groups as well. Although it is widely used in chemical synthesis because no suitable substitutes are available, methylamine is definitely not compatible with the practice of green chemistry. That is because it is highly flammable and toxic. It is a severe irritant to skin, eyes, and mucous membranes of the respiratory tract. It has a noxious odor and is a significant contributor to the odor of rotten fish. In keeping with the reputation of amines as generally unpleasant compounds, another amine, putrescine, gives decayed flesh its characteristic odor.

Many organonitrogen compounds contain oxygen as well. One such compound is nitromethane



Nitromethane

used in chemical synthesis and as a fuel in some race cars. As seen in the structural formula above, the nitro group, $-NO_2$, is the functional group in this compound and related nitro compounds. Another class of organonitrogen compounds also containing oxygen consists of the **nitrosamines**, or N-nitroso compounds, which have figured prominently in the history of green chemistry before it was defined as such. These are compounds that have the N-N=O functional group, which are of concern because several are known carcinogens (cancer-causing agents). The most well known of these is **dimethylnitrosamine** shown below:



Dimethylnitrosamine

This compound used to be employed as an industrial solvent and was used in cutting oils. However, workers exposed to it suffered liver damage and developed jaundice, and the compound as well as other nitrosamines was found to be a carcinogen. A number of other nitrosamines were later found in industrial materials and as byproducts of food processing and preservation. Because of their potential as carcinogens, nitrosamines are avoided in the practice of green chemistry.

Organohalide Compounds

Organohalides exemplified by those shown in Figure 6.5 are organic compounds that contain halogens — F, Cl, Br, or I — but usually chlorine, on alkane, alkene, or aromatic molecules. Organohalides have been widely produced and distributed for a variety of applications, including industrial solvents, chemical intermediates, coolant fluids, pesticides, and other applications. They are for the most part environmentally persistent and, because of their tendency to accumulate in adipose (fat) tissue, they tend to undergo bioaccumulation and biomagnification in organisms

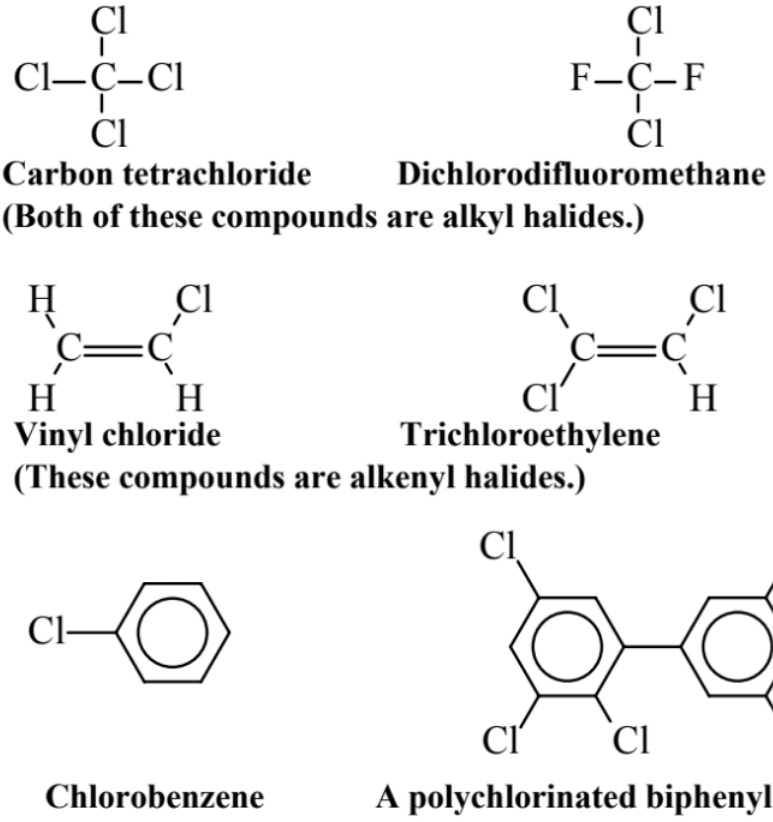


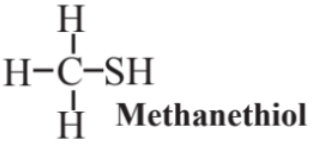
Figure 6.5. Examples of important organohalide compounds including alkyl halides based upon alkanes, alkenyl halides based upon alkenes, and aromatic halides

Carbon tetrachloride is produced when all four H atoms on methane, CH_4, are substituted by Cl. This compound was once widely used and was even sold to the public as a solvent to remove stains and in fire extinguishers, where the heavy CCl_4 vapor smothers fires. It was subsequently found to be very toxic, causing severe liver damage, and its uses are severely restricted. **Dichlorodifluoromethane**is a prominent member of the **chlorofluorocarbon** class of compounds, popularly known as Freons. Developed as refrigerant fluids, these compounds are notably unreactive and nontoxic. However, as discussed in Chapter 10, they were found to be indestructible in the lower atmosphere, persisting to very high altitudes in the stratosphere where chlorine split from them by ultraviolet radiation destroys stratospheric ozone. So the manufacture of chlorofluorocarbons is now prohibited. **Vinyl chloride**, an alkene-based organohalide compound, is widely used to make polyvinylchloride polymers and pipe. Unfortunately, it is a known human carcinogen, so human exposure to it is severely limited. **Trichloroethylene** is an excellent organic solvent that is nonflammable. It is used as a dry cleaning solvent and for degreasing manufactured parts, and was formerly used for food extraction, particularly to decaffeinate coffee. **Chlorobenzene** is the simplest aromatic organochloride. In addition to its uses in making other chemicals, it serves as a solvent and as a fluid for heat transfer. It is extremely stable, and its destruction is a common test for the effectiveness of hazardous waste incinerators. The **polychlorinated biphenyl**(**PCB**) compound shown is one of 209 PCB compounds that can be formed by substituting from 1 to 10 Cl atoms onto the basic biphenyl (two-benzene-ring) carbon skeleton. These compounds are notably stable and persistent, leading to their uses in electrical equipment, particularly as coolants in transformers and in industrial capacitors, as hydraulic fluids, and other applications. Their extreme environmental persistence has led to their being banned. Sediments in New York's Hudson River are badly contaminated with PCBs that were (at the time, legally) dumped or leaked into the river from electrical equipment manufacture from the 1950s into the 1970s.

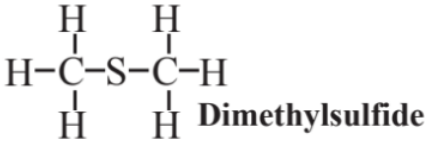
From the discussion above, it is obvious that many organohalide compounds are definitely not green because of their persistence and biological effects. A lot of the effort in the development of green chemistry has been devoted to finding substitutes for organohalide compounds. A 2001 United Nations treaty formulated by approximately 90 nations in Stockholm, Sweden, designated a "dirty dozen" of 12 organohalide compounds of special concern as persistent organic pollutants(POP); other compounds have subsequently been added to this list.

Organosulfur and Organophosphorus Compounds

A number of organosulfur and organophosphorus compounds have been synthesized for various purposes including pesticidal applications. A common class of organosulfur compounds consists of the thiols, the simplest of which is methanethiol:



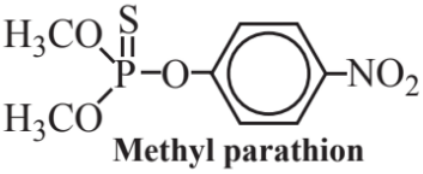
Methanethiol



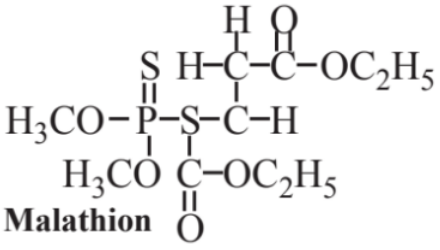
Dimethylsulfide

As with other thiols, which contain the -SH group, this compound is noted for its foul odor. Thiols are added to natural gas so that their odor can warn of gas leaks. Dimethylsulfide, also shown above, is a volatile compound released by ocean-dwelling microorganisms to the atmosphere in such quantities that it constitutes the largest flux of sulfur-containing vapors from Earth to the atmosphere.

Among the most prominent organophosphorus compounds are the organophosphates as shown by methyl parathion and malathion (below). These compounds are both insecticides and contain sulfur as well as phosphorus. Parathion was developed during the 1940s and was once widely used as an insecticide in place of DDT because parathion is very biodegradable, whereas DDT is not and undergoes bioaccumulation and biomagnification in ecosystems. Unfortunately, parathion has a high toxicity to humans and other animals and some human fatalities have resulted from exposure to it. Like other organophosphates, it inhibits acetylcholinesterase, an enzyme essential for nerve function (the same mode of action as its deadly cousins, the “nerve gas” military poisons, such as Sarin). Because of its toxicity, parathion is now banned from general use. Malathion is used in its place and is only about 1/100 as toxic as parathion to mammals because they — though not insects — have enzyme systems that can break it down.



Methyl parathion



Malathion

This page titled 2.5.6.4: Functional Groups is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **6.4: Functional Groups** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.6.5: Giant Molecules from Small Organic Molecules

Reaction 6.2.4 showed the bonding together of molecules of ethylene to form larger molecules. This process, widely practiced in the chemical and petrochemical industries, is called *polymerization* and the products are *polymers*. Many other unsaturated molecules, usually based upon alkenes, undergo polymerization to produce synthetic polymers used as plastics, rubber, and fabrics. As an example, tetrafluoroethylene polymerizes as shown in Figure 6.6 to produce a polymer (Teflon) that is exceptionally heat- and chemical-resistant and that can be used to form coatings to which other materials will not stick (for example, frying pan surfaces).

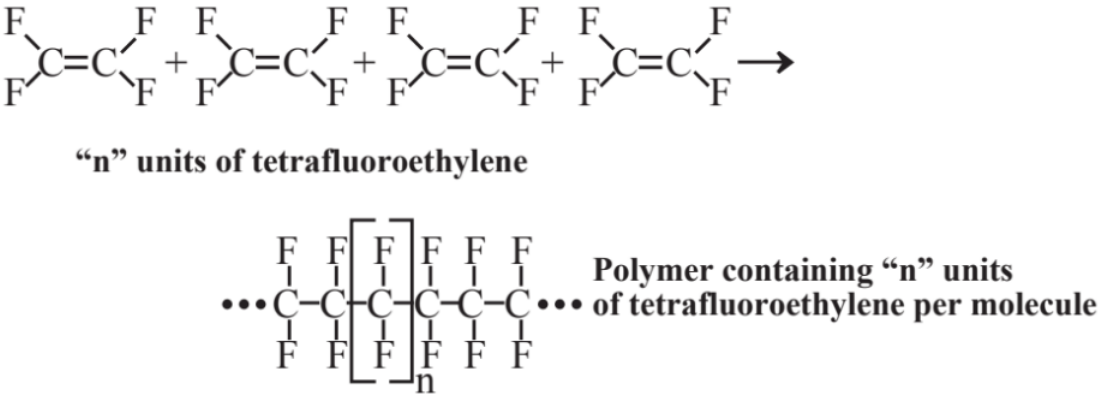
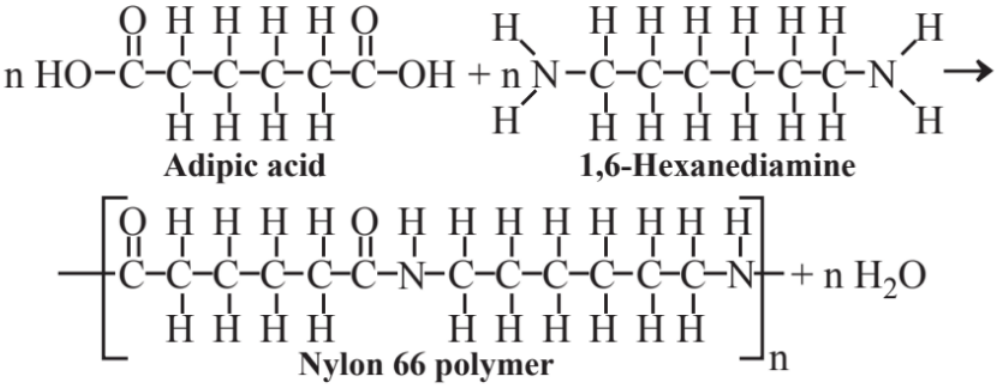


Figure 6.6.
Polymerization of tetrafluoroethylene to produce large molecules of a polymer commonly known as Teflon

Polyethylene and polytetrafluoroethylene are both **additionpolymers** in that they are formed by the chemical addition together of the monomers making up the large polymer molecules. Other polymers are **condensationpolymers** that join together with the elimination of a molecule of water for each monomer unit joined. A common condensation polymer is **nylon**, which is formed by the bonding together of two different kinds of molecules. There are several forms of nylon, the original form of which is nylon 66 discovered by Wallace Carothers, a DuPont chemist, in 1937 and made by the polymerization of adipic acid and 1,6-hexanediamine:

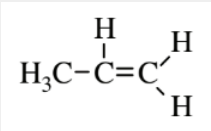
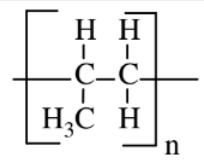
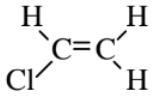
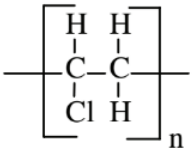
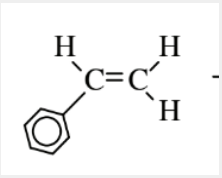
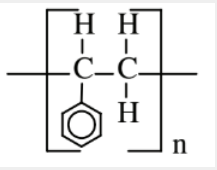
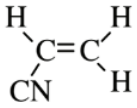
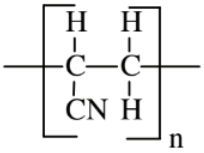
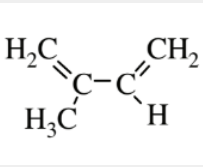
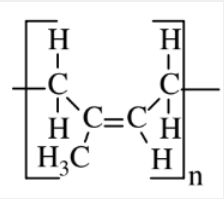


(6.5.1)

There are many different kinds of synthetic polymers that are used for a variety of purposes. Some examples in addition to the ones already discussed in this chapter are given in Table 6.1.

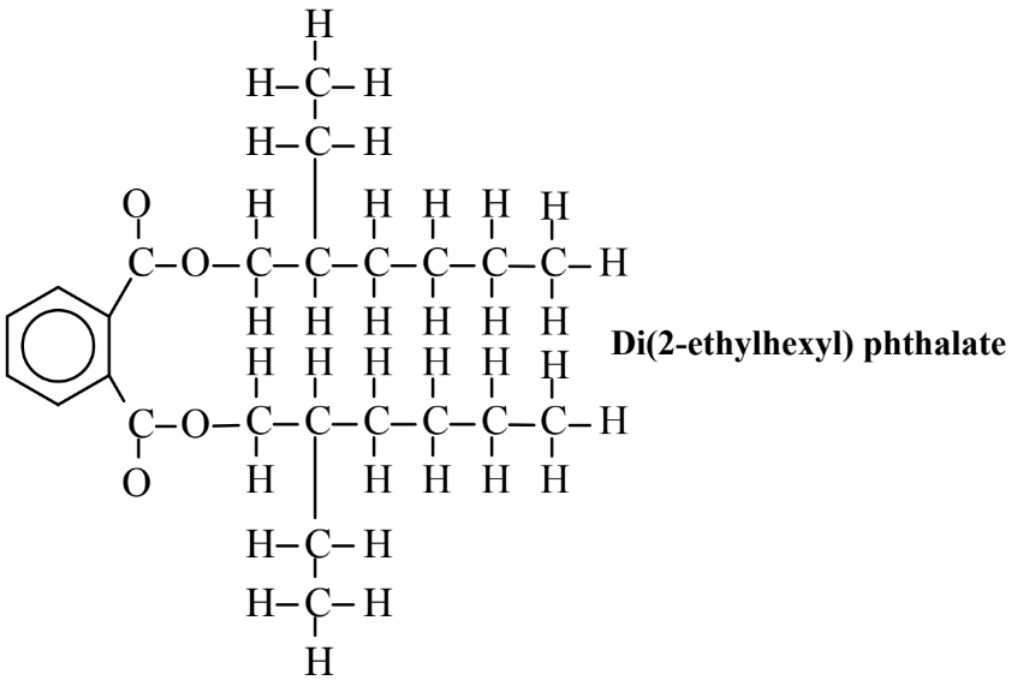
Polymers and the industries upon which they are based are of particular concern in the practice of green chemistry for a number of reasons. The foremost of these is because of the huge quantities of materials consumed in the manufacture of polymers. In addition to the enormous quantities of ethylene and propylene previously cited in this chapter, the U.S. processes about 1.5 billion kg of acrylonitrile, 5.4 billion kg of styrene, 2.0 billion kg of butadiene, and 1.9 kg of adipic acid (for nylon 66) each year to make polymers containing these monomers. These and similarly large quantities of monomers used to make other polymers place significant demands upon petroleum resources and the energy, materials, and facilities required to make the monomers.

Table 6.1. Some Typical Polymers and the Monomers from Which they Are Formed

Monomer	Monomer Formula	Polymer	Applications
Propylene (polypropylene)			Applications requiring harder plastic, luggage, bottles, outdoor carpet
Vinyl chloride (polyvinyl chloride)			Thin plastic wrap, hose, flooring, PVC pipe
Styrene (polystyrene)			Plastic furniture, plastic cups and dishes, blown to produce styrofoam plastic products
Acrylonitrile (polyacrylonitrile)			Synthetic fabrics (Orlon, Acrilan, Creslan), acrylic paints
Isoprene (polyisoprene)			Natural rubber

There is a significant potential for the production of pollutants and wastes from monomer processing and polymer manufacture. Some of the materials contained in documented hazardous waste sites are byproducts of polymer manufacture. Monomers are generally volatile organic compounds with a tendency to evaporate into the atmosphere, and this characteristic combined with the presence of reactive C=C bonds tends to make monomer emissions active in the formation of photochemical smog (see Chapter 10). Polymers, including plastics and rubber, pose problems for waste disposal, as well as opportunities and challenges for recycling. On the positive side, improved polymers can provide long-lasting materials that reduce material use and have special applications, such as liners in waste disposal sites that prevent waste leachate migration and liners in lagoons and ditches that prevent water loss. Strong, lightweight polymers are key components of the blades and other structural components of huge wind generators that are making an increased contribution to renewable energy supplies around the world (see Chapter 16).

Some of the environmental and toxicological problems with polymers have arisen from the use of additives to improve polymer performance and durability. The most notable of these are **plasticizers**, normally blended with plastics to improve flexibility, such as to give polyvinylchloride the flexible characteristics of leather. The plasticizers are not chemically bound as part of the polymer and they leak from the polymer over a period of time, which can result in human exposure and environmental contamination. The most widely used plasticizers are phthalates, esters of phthalic acid as shown by the example of di(2-ethylhexyl) phthalate below. Though not particularly toxic, these compounds are environmentally persistent, resistant to treatment processes, and prone to undergo bioaccumulation. They are found throughout the environment and have been implicated by some toxicologists as possible estrogenic agents that mimic the action of female sex hormone and cause premature sexual development in young female children.



This page titled 2.5.6.5: Giant Molecules from Small Organic Molecules is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **6.5: Giant Molecules from Small Organic Molecules** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.6.6: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. What are two major reactions of alkanes?

2. What is the difference between molecular formulas and structural formulas of organic compounds?

3. What is the difference between ethane and the ethyl group?

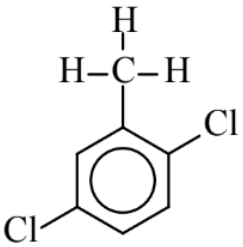
4. What is the structural formula of 3-ethyl-2,3-dimethylpentane?

5. What is a type of reaction that is possible with alkenes, but not with alkanes?

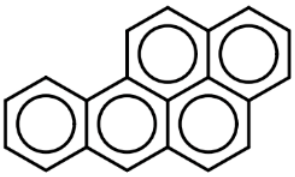
6. What is represented by the structure below?



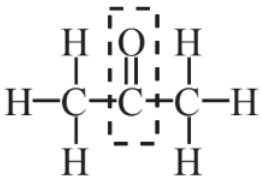
7. Suggest a name for the compound below, which is derived from the hydrocarbon toluene



8. What is a health concern with the aromatic compound below:



9. What do the groups of atoms outlined by dashed lines represent in the structure below?



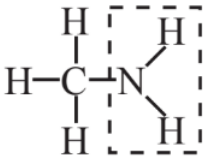


Figure 6.4, what are the similarities and differences between organic oxides and ethers.

11. What are 3 separate kinds of groups characteristic of organonitrogen compounds?

12. What is a class of organochlorine compounds consisting of many different kinds of molecules that is noted for environmental persistence?

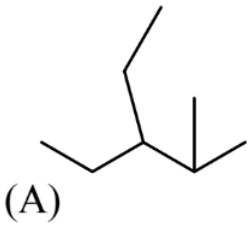
13. What is a notable characteristic of organosulfur thiols?

14. What is a particularly toxic organophosphorus compound? What is a biochemical molecule containing phosphorus?

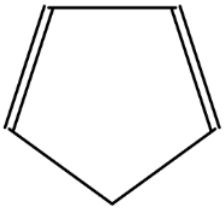
15. What are polymers and why are they important?

16. Examination of the formulas of many of the monomers used to make polymers reveals a common characteristic. What is this characteristic and how does it enable polymer formation? Does nylon illustrate a different pathway to monomer formation? Explain.

17. Write the complete structural formulas corresponding to each of the line structures below:



(A)



(B)

18. Some of the most troublesome organic pollutant compounds are organochlorine compounds including the “dirty dozen” persistent organic pollutants mentioned in this chapter. Organohalides involving a halogen other than chlorine are emerging as significant pollutants. Doing some research on the internet, find which class of compounds these are and why they are significant pollutants.

2.5.6.6: Questions and Problems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** is licensed CC BY-NC-SA 4.0.

2.5.6.7: Supplementary References

Armstrong, James, *General, Organic, and Biochemistry*, Brooks/Cole, Cengage Learning, Belmont, CA, 2010.

Bettelheim, Frederick A., *Introduction to General, Organic, and Biochemistry*, 9th ed., Brooks/Cole, Cengage Learning, Belmont, CA, 2010.

Denniston, Katherine J., Joseph J. Topping, and Robert L. Caret, *General, Organic, and Biochemistry*, 7th ed., McGraw-Hill, New York, 2011.

Guinn, Denise, and Rebecca Brewer, *Essentials of General, Organic, and Biochemistry: An Integrated Approach*, W. H. Freeman and Co., New York, 2009.

McMurry, John, David S. Ballantine, Carl A. Hoeger, Virginia E. Peterson, and Mary E.Castellion, *Fundamentals of General, Organic, and Biological Chemistry*, 6th ed., Prentice-Hall, Upper Saddle River, NJ, 2009.

Seager, Spencer L., and Michael R. Slabaugh, *Organic and Biochemistry for Today*, 7th ed., Brooks/Cole, Cengage Learning, Belmont, CA, 2010.

Solomons, T.W. Graham, and Craig Fryhle, *Organic Chemistry*, 10th ed., Wiley, Hoboken, NJ, 2011.

Smith, Janice G., *Principles of General, Organic, and Biochemistry*, McGraw-Hill, New York, 2011.

Stoker, H. Stephen, *General, Organic, and Biological Chemistry*, 5th ed., Brooks/Cole, Cengage Learning, Belmont, CA, 2010.

Winter, Arthur, *Organic Chemistry 1 for Dummies*, Wiley, Hoboken, NJ, 2005.

2.5.6.7: Supplementary References is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Supplementary References** is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

2.5.7: Chemistry of Life and Green Chemistry

“A microscopic cell of photosynthetic cyanobacteria constitutes a complex of chemical factories that carry out a multitude of biochemical processes. Powered by solar energy and operating under ambient conditions, these organisms take carbon dioxide and nitrogen from air and simple inorganic ions dissolved in water and make all the life molecules they need for their metabolism and reproduction. In eons past these kinds of organisms generated all of the oxygen that is in Earth’s atmosphere. For all their knowledge of chemistry it would be impossible for humans to reproduce the chemical processes of these remarkable bacteria”

2.5.7.1: Green Biochemistry

2.5.7.2: Biochemistry and the Cells

2.5.7.3: Carbohydrates

2.5.7.4: Proteins

2.5.7.5: Lipids - Fats, Oils, and Hormones

2.5.7.6: Nucleic Acids

2.5.7.7: Enzymes

2.5.7.8: Biochemical Processes in Metabolism

2.5.7.9: Biochemistry of Toxic Substances and Toxicological Chemistry

2.5.7.10: Literature Cited

2.5.7.11: Questions and Problems

2.5.7.12: Supplementary References

This page titled 2.5.7: Chemistry of Life and Green Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.7.1: Green Biochemistry

Biochemistry is the science of chemical processes that occur in living organisms.[1] By its nature biochemistry is a green chemical and biological science. This is because over eons of evolution organisms have evolved that carry out biochemical processes sustainably. Because the enzymes that carry out biochemical processes can only function under mild conditions, particularly of temperature, biochemical processes take place under safe conditions, avoiding the high temperatures, high pressures, and corrosive and reactive chemicals that often characterize synthetic chemical operations. Therefore, it is appropriate to refer to **green biochemistry**.

The ability of organisms to carry out chemical processes is truly amazing, even more so when one considers that many of them occur in single-celled organisms. Photosynthetic cyanobacteria consisting of individual cells less than a micrometer (μm) in size can make all the complex biochemicals they need to exist and reproduce using sunlight for energy and simple inorganic substances such as CO_2, K^+ ion, NO_3^- ion and HPO_4^{2-} ion for raw materials. Beginning soon after conditions on Earth became hospitable to life, these photosynthetic bacteria produced the oxygen that now composes about 20% of Earth's atmosphere. Fossilized stromatolites (bodies of sedimentary materials bound together by films produced by microorganisms) produced by cyanobacteria have been demonstrated dating back 2.8 billion years, and this remarkable microorganism that converts atmospheric carbon dioxide to biomass and atmospheric N_2 to chemically fixed N may have been on Earth as long as 3.5 billion years ago.

It is fascinating to view single live cells of animal-like protozoa through an optical microscope. An ameba appears as a body of cellular protoplasm and moves by oozing about like a living blob of jelly. Examination of *Euglena*protozoa may show a cell several μm in size with many features including a cell nucleus that serves to direct metabolism and reproduction, green chloroplasts for photosynthetic production of biomass, a red eye-spot sensitive to light, a mouth-like contractile vacuole by which the cell expels excess water, and a thin tail-like structure(flagella) that moves rapidly and propels the cell. More detailed examination by electron microscope of such cells and those that make up more complex organisms reveals many more cell parts that are involved with biochemical function.

At least a rudimentary knowledge of biochemistry is needed to understand green chemistry, environmental chemistry, and sustainability science and technology. One reason is the ability of organisms to synthesize a vast variety of substances. The most obvious of these is biomass made by the photosynthetic fixation of carbon dioxide and that forms the basis of nature's food webs.Organisms make many of the materials upon which humans rely. In addition to food, one such material is the lignocellulose that composes most of plant biomass such as wood used for construction, paper-making, and fuel. Very complex molecules are made by organisms, for example, human insulin produced by genetically engineered organisms. Organisms make materials under very mild conditions compared to those used in the anthrosphere. An important example is chemically fixed nitrogen from the atmosphere which is produced synthetically in the anthrosphere as ammonia (NH_3) at high temperatures and pressures whereas *Rhizobium* bacteria attached to the roots of soybeans and other legume plants fix nitrogen in the mild conditions of the soil environment. Increasingly as supplies of petroleum and other non-renewable raw materials become more scarce, humans are turning to microorganisms and plants to make essential materials.

Another major reason for considering biochemistry as part of green chemistry and sustainability is the protection of organisms from products and processes in the anthrosphere. It is essential to know the potential toxic effects of various materials, a subject addressed by **toxicological chemistry**.[2] One of the fundamental goals of green chemistry is to minimize the production and use of products that may have adverse environmental effects. Sustainability of the entire planet requires that humans not disperse into the environment substances that may undergo bioaccumulation and be toxic to humans and other organisms.

Biochemical processes not only are profoundly influenced by chemical species in the environment, they largely determine the nature of these species, their degradation, and even their syntheses, particularly in the aquatic and soil environments. The study of such phenomena forms the basis of **environmental biochemistry.**

This chapter is designed to give an overview of biochemistry and how it relates to green chemistry and sustainability science and technology. A glance at the structural formulas of some of the biochemicals shown in this chapter gives a hint of the complexity of biochemistry. This chapter is designed to provide a basic understanding of this complex science with enough detail for it to be meaningful but to avoid overwhelming the reader. It begins with an overview of the four major classes of biochemicals— proteins, carbohydrates, lipids, and nucleic acids. Many of the compounds in these classes are polymers with molecular masses of the order

of a million or even larger. Proteins and nucleic acids consist of macromolecules, lipids are usually relatively small molecules, carbohydrates range from small sugar molecules to high-molar-mass macromolecules such as those in cellulose.

The behavior of a substance in a biological system depends to a large extent upon whether the substance is hydrophilic (“water-loving”) or hydrophobic (“water-hating”). Some important toxic substances are hydrophobic, a characteristic that enables them to traverse cell membranes readily and to bioaccumulate in lipid (fat) tissue. Many hydrocarbons and organohalide compounds synthesized from hydrocarbons are hydrophobic. Part of the detoxification process carried on by living organisms is to render such molecules hydrophilic, therefore water-soluble and readily eliminated from the body.

This page titled 2.5.7.1: Green Biochemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **7.1: Green Biochemistry** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.7.2: Biochemistry and the Cells

For the most part, biochemical processes occur within cells, the very small units of which living organism are composed.[3] Cells are discussed in more detail as basic units of life in Chapter 12, Section 12.3; here they are regarded as what chemical engineers would call “unit operations” for carrying out biochemical processes. Many organisms consist of single cells or individual cells growing together in colonies. Bacteria, yeasts, protozoa, and some algae consist of single cells. Other than these microorganisms, organisms are composed of many cells that have different functions. Liver cells, muscle cells, brain cells, and skin cells in the human body are quite different from each other and do different things. Two major kinds of cells are **eukaryotic** cells which have a nucleus and **prokaryotic** cells which do not. Prokaryotic cells are found predominately in single-celled bacteria. Eukaryotic cells occur in multicellular plants and animals — higher life forms.

Cell structure has an important influence on determining the nature of biomaterials. Muscle cells consist largely of strong structural proteins capable of contracting and movement. Bone cells secrete a protein mixture that then mineralizes with calcium and phosphate to produce solid bone.The walls of cells in plants are largely composed of strong cellulose, which makes up the sturdy structure of wood.

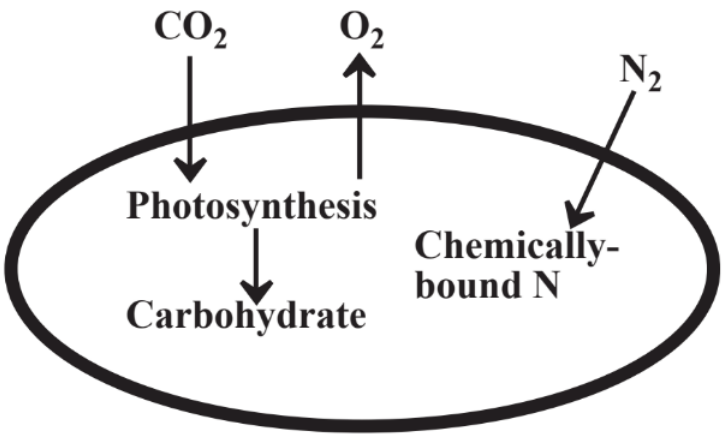


Figure 7.1. Cyanobacteria are remarkable organisms that within a single “simple” prokaryotic cell carry out all the biochemical processes needed to convert atmospheric carbon dioxide to carbohydrate and biomass and that can split the chemically very stable atmospheric nitrogen molecule and convert the nitrogen to chemically and biochemically bound nitrogen

This page titled 2.5.7.2: Biochemistry and the Cells is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **7.2: Biochemistry and the Cells** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.7.3: Carbohydrates

Carbohydrates are biomolecules consisting of carbon, hydrogen, and oxygen having the approximate simple formula CH_2O. One of the most common carbohydrates is the simple sugar glucose shown in Figure 7.2. Units of glucose and other simple sugars called **monosaccharides** join together in chains with the loss of a water molecule for each linkage to produce macromolecular **polysaccharides**. These include **starch** and **cellulose** in plants and starch-like glycogen in animals.

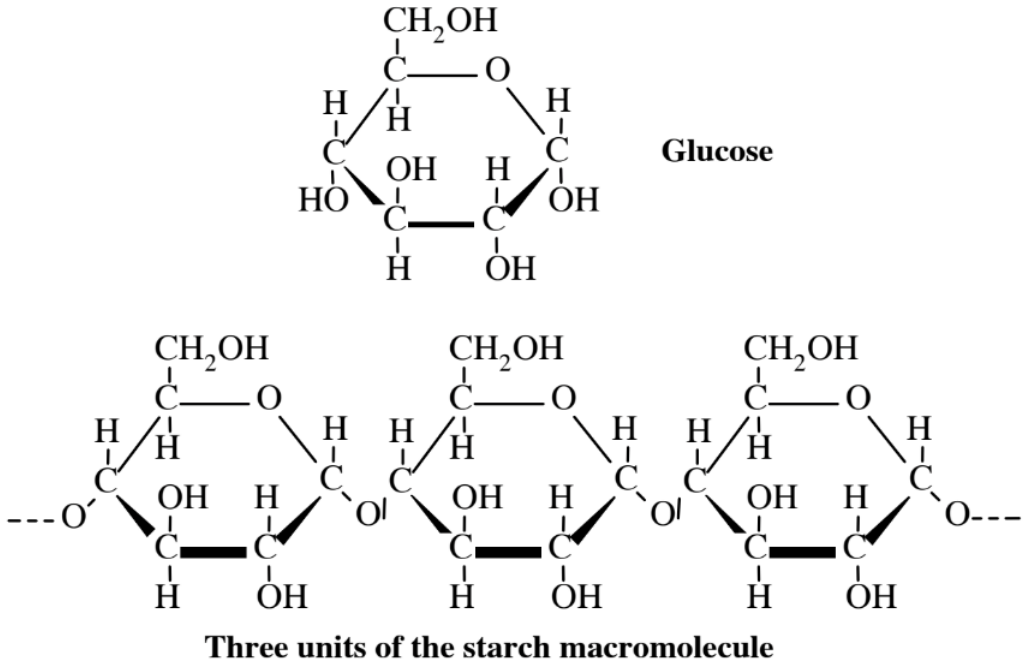


Figure 7.2. Glucose, a monosaccharide, or simple sugar, and a segment of the starch molecule, which is formed when glucose molecules polymerize with the elimination of one H_2O molecule per glucose monomer.

Glucose carbohydrate is the biological material generated from water and carbon dioxide when solar energy in sunlight is utilized in photosynthesis. The overall reaction is

$$6\,CO_2 + 6\,H_2O \rightarrow C_6H_{12}O_6 + 6\,O_2$$

This is obviously an extremely important reaction because it is the one by which inorganic molecules are used to synthesize high-energy carbohydrate molecules that are in turn converted to the vast number of biomolecules that comprise living systems. There are other simple sugars, including fructose, mannose, and galactose, that have the same simple formula as glucose, $C_6H_{12}O_6$, but which must be converted to glucose before being utilized by organisms for energy. Common table sugar, sucrose, $C_{12}H_{22}O_{11}$, consists of a molecule of glucose and one of fructose linked together (with the loss of a water molecule); because it is composed of two simple sugars sucrose is called a **disaccharide**.

Starch molecules, which may consist of several hundred glucose units joined together, are readily broken down by organisms to produce simple sugars used for energy and to produce biomass. For example, humans readily digest starch in potatoes or bread to produce glucose used for energy (or to make fat tissue).

The chemical formula of starch is $(C_6H_{10}O_5)_n$, where *n* may represent a number as high as several hundreds. What this means is that the very large starch molecule consists of as many as several hundred units of $C_6H_{10}O_5$ from glucose joined together. For example, if *n* is 100, there are 6 times 100 carbon atoms, 10 times 100 hydrogen atoms, and 5 times 100 oxygen atoms in the molecule. Its chemical formula is $C_{600}H_{1000}O_{500}$. The atoms in a starch molecule are actually present as linked rings represented by the structure shown in Figure 7.2. Starch occurs in many foods, such as bread, potatoes, and cereals. It is readily digested by animals, including humans.

Cellulose is a polysaccharide which is also made up of $C_6H_{10}O_5$ units. Molecules of cellulose are huge, with molecular masses of around 400,000. The cellulose structure (Figure 7.3) is similar to that of starch. Cellulose is produced by plants and forms the

structural material of plant cell walls. Wood is about 60% cellulose, and cotton contains over 90% of this material. Fibers of cellulose are extracted from wood and pressed together to make paper.

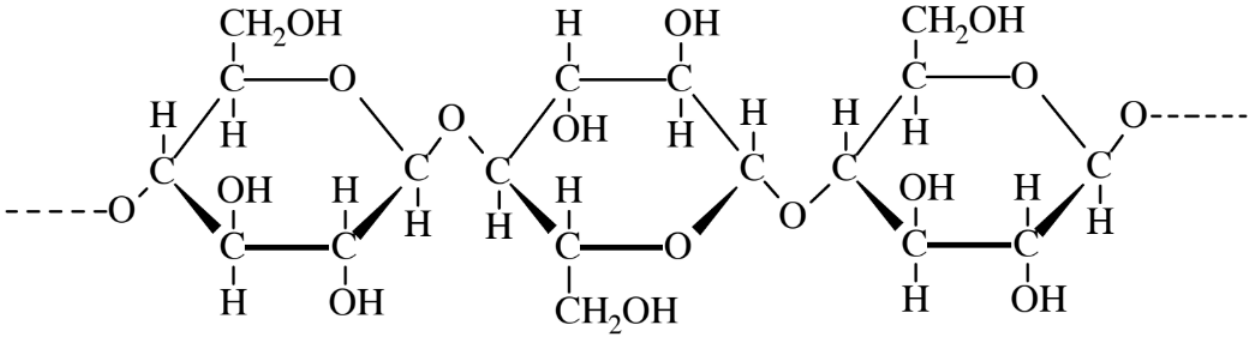


Figure 7.3. A segment of a cellulose molecule. These molecules are biosynthesized from glucose with the loss of one H_2O for each linkage formed

Humans and most other animals cannot digest cellulose because they lack the enzyme needed to hydrolyze the oxygen linkages between the glucose molecules. Ruminant animals (cattle, sheep, goats, moose) have bacteria in their stomachs that break down cellulose into products which can be used by the animal. Fungi and termites existing synergistically with cellulose-degrading bacteria biodegrade huge quantities of cellulose. Chemical processes are available to convert cellulose to simple sugars by the reaction

$$\underbrace{(C_6H_{10}O_5)_n}_{\text{cellulose}} + n\,H_2O \rightarrow \underbrace{n\,C_6H_{12}O_6}_{\text{glucose}}$$

where *n* may be 2000-3000. This involves breaking the linkages between units of $C_6H_{10}O_5$ by adding a molecule of H_2O at each linkage, a hydrolysis reaction. Large amounts of cellulose from wood, sugar cane, and agricultural products go to waste each year. The hydrolysis of cellulose enables these products to be converted to sugars, which can be fed to animals. The potential for producing very large quantities of glucose from cellulose have led to intense efforts to hydrolyze cellulose to glucose with enzymes (biological catalysts) and is an important effort in green chemistry.

Carbohydrates are potentially very important in green chemistry. For one thing, they are a concentrated form of organic energy synthesized and stored by plants as part of the process by which plants capture solar energy through photosynthesis. Carbohydrates can be utilized directly for energy or fermented to produce ethanol, C_2H_6O, a combustible alcohol that is added to gasoline or can even be used in place of gasoline. Secondly, carbohydrates are a source of organic raw material that can be converted to other organic molecules to make plastics and other useful materials.

This page titled 2.5.7.3: Carbohydrates is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **7.3: Carbohydrates** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.7.4: Proteins

Proteins are macromolecules that are composed of nitrogen, carbon, hydrogen, and oxygen along with smaller quantities of sulfur. The small molecules of which proteins are made are composed of 20 naturally occurring **amino acids**. The simplest of these, glycine, is shown in the first structure in Figure 7.4, along with two other amino acids. As shown in Figure 7.4, amino acids join together with the loss of a molecule of H_2O for each linkage formed. The three amino acids in Figure 7.4 are shown linked together as they would be in a protein in the bottom structure in the figure. Many hundreds of amino acid units may be present in a protein molecule.

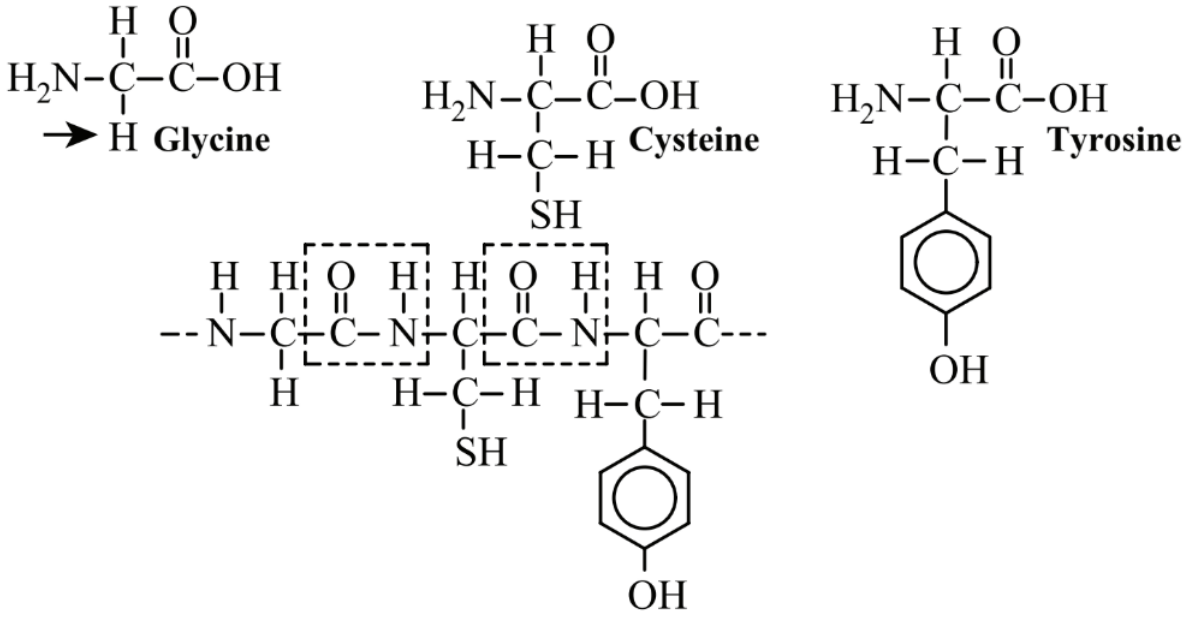


Figure 7.4.
Three amino acids. Glycine is the simplest amino acid. All others have the basic glycine structure except that different groups are substituted for the H designated in glycine by an arrow. The lower structure shows these three amino acids linked together in a macromolecule chain composing a protein. For each linkage, one molecule of H_2O is lost. The peptide linkages holding amino acids together in proteins is outlined by a dashed rectangle.

The three-dimensional structures of protein molecules are of the utmost importance and largely determine what the proteins do in living systems and how they are recognized by other biomolecules. Enzymes, special proteins that act as catalysts to enable biochemical reactions to occur, recognize the substrates upon which they act by the complementary shapes of the enzyme molecules and substrate molecule. There are several levels of protein structure. The first of these is determined by the order of amino acids in the protein macromolecule. Folding of protein molecules and pairing of two different protein molecules further determine structure. The loss of protein structure, called **denaturation**, can be very damaging to proteins and to the organism in which they are contained.

Two major kinds of proteins are tough **fibrousproteins** that compose hair, tendons, muscles, feathers, and silk, and spherical or oblong-shaped **globular proteins**, such as hemoglobin in blood or the proteins that comprise enzymes. Proteins serve many functions. These include **nutrient proteins**, such as casein in milk, **structural proteins**, such as collagen in tendons, **contractile proteins**, such as those in muscle, and **regulatory proteins**, such as insulin, that regulate biochemical processes.

Proteins with carbohydrate groups attached constitute an important kind of biomolecule called **glycoproteins**. Collagen is a crucial glycoprotein that provides structural integrity to body parts. It is a major constituent of skin, bones, tendons, and cartilage.

Some proteins are very valuable biomaterials for pharmaceutical, nutritional, and other applications and their synthesis is an important aspect of green chemistry. The production of specific proteins has been greatly facilitated in recent years by the application of genetic engineering to transfer to bacteria the genes that direct the synthesis of specific proteins. The best example is insulin, a protein injected into diabetics to control blood sugar. Insulin injected for blood glucose control used to be isolated from the pancreas's of slaughtered cattle and hogs. Although this enabled many diabetics to live normal lives, the process of getting the

insulin was cumbersome, supply was limited, and the insulin from this source was not exactly the same as that made in the human body, which often caused the body to have an allergic response to it as a foreign protein. The transfer through recombinant DNA technology of the human gene for insulin production into prolific *Escherichiacoli* bacteria has enabled large-scale production of human insulin by the bacteria.

This page titled 2.5.7.4: Proteins is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **7.4: Proteins** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.7.5: Lipids - Fats, Oils, and Hormones

Lipids differ from most other kinds of biomolecules in that they are repelled by water. Lipids can be extracted from biological matter by organic solvents, such as diethyl ether or toluene. Recall that proteins and carbohydrates are distinguished largely by chemically similar characteristics and structures. However, lipids have a variety of chemical structures that share the common physical characteristic of solubility in organic solvents. Many of the commonly encountered lipid fats and oils are esters of glycerol alcohol, $CH_2(OH)CH(OH)CH_2(OH)$, and long-chain carboxylic acids (fatty acids), such as stearic acid, $CH_3(CH_2)_{16}CO_2H$. The glycerol molecule has three -OH groups to each of which a fatty acid molecule may be joined through the carboxylic acid group with the loss of a water molecule for each linkage that is formed. Figure 7.5shows a fat molecule formed from three stearic acid molecules and a glycerol molecule. Such a molecule is one of many possible **triglycerides**. Also shown in this figure is cetyl palmitate, the major ingredient of spermaceti wax extracted from sperm whale blubber and used in some cosmetics and pharmaceutical preparations. Cholesterol shown in Figure 7.5 is one of several important lipid **steroids**, which share the ring structure composed of rings of 5 and 6 carbon atoms shown in the figure for cholesterol.

Although the structures shown in Figure 7.5 are diverse, they all share a common characteristic. This similarity is the preponderance of hydrocarbon chains and rings so that lipid molecules largely resemble hydrocarbons. Their hydrocarbon-like molecules make lipids soluble in organic solvents.

Some of the steroid lipids are particularly important because they act as **hormones**, chemical messengers that convey information from one part of an organism to another. Major examples of steroid hormones are cholesterol, testosterone (male sex hormone), and estrogens (female sex hormones). Steroid lipids readily penetrate the membranes that enclose cells, which are especially permeable to more hydrophobic lipid materials. Hormones, start and stop a number of body functions and regulate the expression of many genes. In addition to steroid lipids, many hormones including insulin and human growth hormone are proteins. Hormones are given off by ductless glands in the body called **endocrine glands**. The locations of important endocrine glands are shown in Figure 7.6.

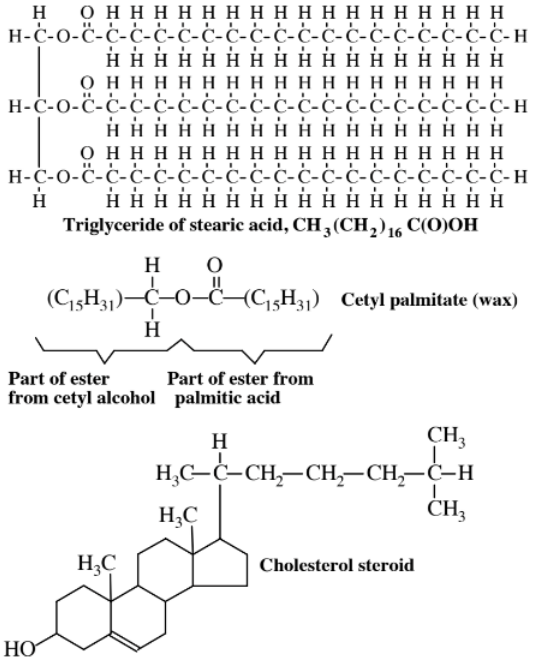


Figure 7.5. Three examples of lipids formed in biological systems. Note that a line structure is used to show the ring structure of cholesterol. The hydrocarbon-like nature of these compounds which makes them soluble in organic compounds is obvious.

Lipids are important in green chemistry for several reasons. Lipids are very much involved with toxic substances, the generation and use of which are always important in green chemistry. Poorly biodegradable substances, particularly organochlorine compounds, that are always an essential consideration in green chemistry, tend to accumulate in lipids in living organisms, a process called bioaccumulation. Lipids can be valuable raw materials and fuels. A major kind of renewable fuel is made by hydrolyzing the long-chain fatty acids from triglycerides and attaching methyl groups to produce esters. This liquid product,

commonly called biodiesel fuel, serves as a substitute for petroleum-derived liquids in diesel engines. The development and cultivation of plants that produce oils and other lipids is a major possible route to the production of renewable resources.

This page titled 2.5.7.5: Lipids - Fats, Oils, and Hormones is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **7.5: Lipids - Fats, Oils, and Hormones** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.7.6: Nucleic Acids

Nucleic acids (Figure 7.7)are biological macromolecules that store and pass on the genetic information that organisms need to reproduce and synthesize proteins. The two major kinds of nucleic acids are **deoxyribonucleic acid**,**DNA**, which basically stays in place in the cell nucleus of an organism and **ribonucleic acid**, **RNA**, which is spun off from DNA and functions throughout a cell. Molecules of nucleic acids contain three basic kinds of materials. The first of these is a simple sugar, 2-deoxy-β-D-ribofuranose (deoxyribose) contained in DNA and β-D-ribofuranose (ribose) contained in RNA. The second major kind of ingredient consists of nitrogen-containing bases: cytosine, adenine, and guanine, which occur in both DNA and RNA, thymine, which occurs only in DNA, and uracil, which occurs only in RNA. The third constituent of both DNA and RNA is inorganic phosphate, PO_4^{3-}. These three kinds of substances occur as repeating units called **nucleotides** joined together in astoundingly long chains in the nucleic acid polymer as shown in Figure 7.7.

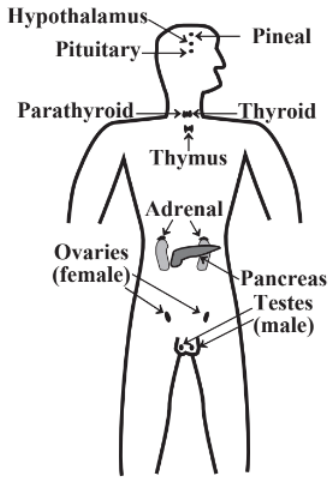


Figure 7.6. Location of important endocrine ductless glands that secrete hormones directly into the bloodstream in the human body. In contrast, exocrine glands including sweat, salivary, and mammary glands secrete hormones through ducts.In addition to secreting insulin into the bloodstream, the pancreas also functions as a ducted gland secreting pancreatic juice into the small intestine. This material contains digestive enzymes that help break down carbohydrate, protein, and fat molecules

The remarkable way in which DNA operates to pass on genetic information and perform other functions essential for life is the result of the structure of the DNA molecule. In 1953, James D.Watson, and Francis Crick deduced that DNA consisted of two strands of material counterwound around each other in a structure known as anα-helix (Figure 7.8), a remarkable bit of insight that earned Watson and Crick the Nobel Prize in 1962. These strands are held together by hydrogen bonds between complementary nitrogenous bases. Taken apart, the two strands resynthesize complementary strands, a process that occurs during reproduction of cells in living organisms. Indirecting protein synthesis, DNA becomes partially unravelled and generates a complementary strand of material in the form of RNA, which in turn directs protein synthesis in the cell.

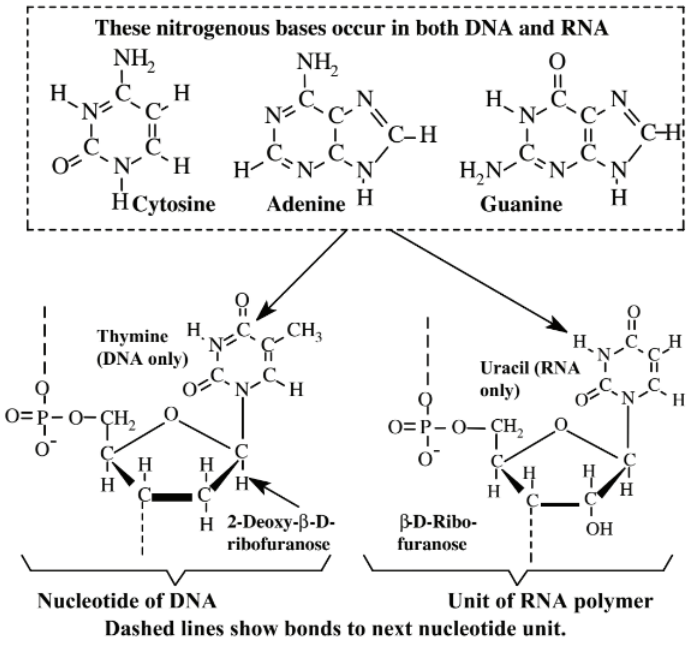


Figure 7.7. Basic units of nucleic acid polymers. These units act as a code in directing reproduction and other activities of organisms.

Consideration of nucleic acids and their function is very important in the development of green chemistry. One aspect of this relationship is that the toxicity hazards of many chemical substances result from potential effects of these substances upon DNA. Of most concern is the ability of some substances to alter DNA and cause uncontrolled cell replication characteristic of cancer. Also of concern is the ability of some chemical substances called **mutagens** to alter DNA such that undesirable characteristics are passed on to offspring.

Another important consideration with DNA as it relates to green chemistry is the ability that humans now have to transfer DNA between organisms, popularly called genetic engineering. An important example is the development of bacteria that have the DNA transferred from humans to make human insulin. This technology of recombinant DNA is discussed in more detail in Chapter 12.

This page titled 2.5.7.6: Nucleic Acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **7.6: Nucleic Acids** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.7.7: Enzymes

Recall from Chapter 5, Section 5.5, that catalysts are substances that speed up a chemical reaction without themselves being consumed in the reaction. Catalysis is one of the most important aspects of green chemistry because the ability to make reactions go faster as well as more efficiently, safely, and specifically means that less energy and raw materials are used and less waste is produced. Biochemical catalysts called **enzymes** include some of the most sophisticated of catalysts. Enzymes speed up biochemical reactions by as much as ten- to a hundred million-fold. They often enable reactions to take place that otherwise would not occur, that is, they tend to be very selective in the reactions they promote. One of the greatest advantages of enzymes as catalysts is that they have evolved to function under the benign conditions under which organisms exist. This optimum temperature range is generally from about the freezing point of water (0°C) to slightly above body temperature (up to about 40°C). Chemical reactions go faster at higher temperatures, so there is considerable interest in enzymes isolated from microorganisms that thrive at temperatures near the boiling point of water (100°C) in hot water pools heated by underground thermal activity such as are found in Yellowstone National Park.

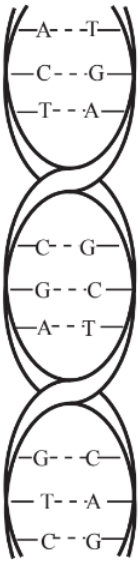


Figure 7.8. Representation of the double helix structure of DNA showing the allowed base pairs held together by hydrogen bonding between the phosphate/sugar polymer "backbones" of the two strands of DNA. The letters stand for adenine (A), cytosine (C), guanine (G), and thymine (T). The dashed lines, ---, represent hydrogen bonds.

Enzymes are proteinaceous substances. Their structure is highly specific so that they bind with whatever they act upon, a substance called a **substrate**. The basic mechanism of enzyme action is shown in Figure 7.9. As indicated by the figure, an enzyme recognizes a substrate by its shape, bonds with the substrate to produce an **enzyme-substrate complex**, causes a change such as splitting a substrate in two with addition of water (hydrolysis), then emerges unchanged to do the same thing again. The basic process can be represented as follows:

$$\text{enzyme} + \text{substrate} \leftrightarrows \text{enzyme-substrate complex} \leftrightarrows \text{enzyme} + \text{product}$$

Note that the arrows in the formula for enzyme reaction point both ways. This means that the reaction is **reversible**. An enzyme-substrate complex can simply go back to the enzyme and the substrate. The products of an enzymatic reaction can react with the enzyme to form the enzyme-substrate complex again. It, in turn, may again form the enzyme and the substrate. Therefore, the same enzyme may act to cause a reaction to go either way.

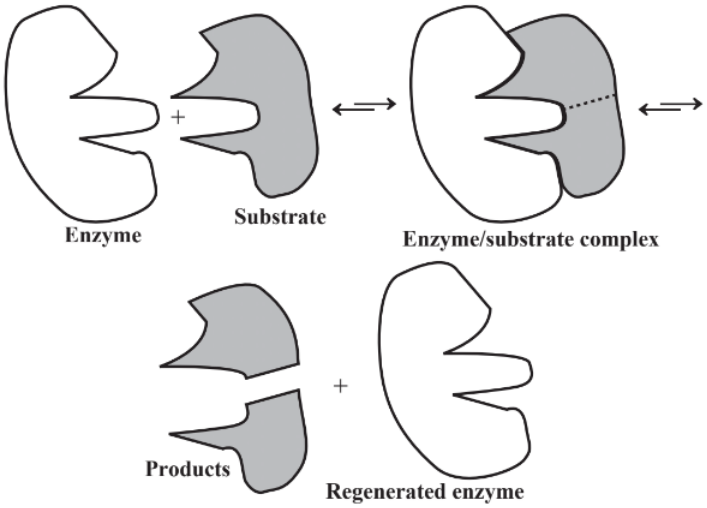


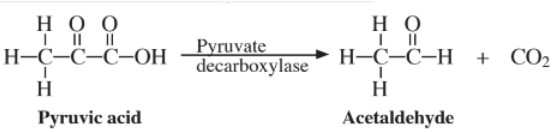
Figure 7.9. Representation of the “lock-and-key” mode of enzyme action which enables the very high specificity of enzyme-catalyzed reactions

In order for some enzymes to work, they must first be attached to **coenzymes**. Coenzymes normally are not protein materials. Some of the vitamins are important coenzymes.

The names of enzymes are based upon what they do and where they occur. For example, *gastric protease*, commonly called pepsin,is an enzyme released by the stomach (gastric), which splits protein molecules as part of the digestion process (protease). Similarly, the enzyme produced by the pancreas that breaks down fats (lipids) is called *pancreatic lipase*. Its common name is steapsin. In general, lipase enzymes cause lipid triglycerides to dissociate and form glycerol and fatty acids.

Lipase and protease enzymes are **hydrolyzing enzymes**, which enable the breakdown of high-molecular-mass biological compounds and add water, one of the most important types of the reactions involved in digestion of food carbohydrates, proteins, and fats. Recall that the higher carbohydrates humans eat are largely disaccharides (sucrose, or table sugar) and polysaccharides(starch). These are formed by the joining together of units of simple sugars, $C_6H_{12}O_6$, with the elimination of an H_2O molecule at the linkage where they join. Proteins are formed by the condensation of amino acids, again with the elimination of a water molecule at each linkage. Fats are esters which are produced when glycerol and fatty acids link together. A water molecule is lost for each of these linkages when a protein, fat, or carbohydrate is synthesized. In order for these substances to be used as a food source, the reverse process must be catalyzed by hydrolyzing enzymes to break down large, complicated molecules of protein, fat, or carbohydrate to simple, soluble substances which can penetrate a cell membrane and take part in chemical processes in the cell.

An important biochemical process is the shortening of carbon atom chains, such as those in fatty acids, commonly by the elimination of CO_2 from carboxylic acids. For example, *pyruvate decarboxylase* enzyme removes CO_2 from pyruvic acid,



to produce a compound with one less carbon. It is by such carbon-by-carbon breakdown reactions that long chain compounds are eventually degraded to CO_2 in the body. Another important consequence of this kind of reaction is the biodegradation of long-chain hydrocarbons by the action of microorganisms in the water and soil environments.

Energy is exchanged in living systems largely by oxidation and reduction mediated by **oxidoreductase enzymes**. Cellular respiration is an oxidation reaction in which a carbohydrate, $C_6H_{12}O_6$, is broken down to carbon dioxide and water with the release of energy

$$C_6H_{12}O_6 + 6\,O_2 \rightarrow 6\,CO_2 + 6\,H_2O + \text{energy}$$

https://bio.libretexts.org/@go/page/108826

Actually, such an overall reaction occurs in living systems by a complicated series of individual steps including oxidation. The enzymes that bring about oxidation in the presence of free O2 are called **oxidases.**

In addition to the major types of enzymes discussed above there are numerous other enzymes that perform various functions. **Isomerases** form isomers of particular compounds. For example, isomerases convert several simple sugars with the formula $C_6H_{12}O6$ to glucose, the only sugar that can be used directly for cell processes. **Transferase enzymes** move chemical groups from one molecule to another, **lyase enzymes** remove chemical groups without hydrolysis and participate in the formation of C=C bonds or addition of species to such bonds, and **ligase enzyme**s work in conjunction with ATP (adenosine triphosphate, a high-energy molecule that plays a crucial role in energy-yielding, glucose-oxidizing metabolic processes) to link molecules together with the formation of bonds such as carbon-carbon or carbon-sulfur bonds.

Enzymes are affected by the conditions and media in which they operate. Among these is the hydrogen ion concentration (pH). An interesting example is gastric protease which requires the acid environment of the stomach to work well but stops working when it passes into the much more alkaline medium of the small intestine. This prevents damage to the intestine walls, which would occur if the enzyme tried to digest them. Part of the damage to the esophagus from reflux esophagitis (acid reflux) is due to the action of gastric protease enzyme that flows back into the esophagus from the stomach with the acidic stomach juices. Temperature is critical for enzyme function. Not surprisingly, the enzymes in the human body work best at around 37° C (98.6° F), which is the normal body temperature. Heating these enzymes to around 60° C permanently destroys them. Some bacteria that thrive in hot springs have enzymes that work best at temperatures as high as that of boiling water. Other “cold-seeking” bacteria have enzymes adapted to near the freezing point of water.

Immobilized Enzymes in Green Chemistry

As noted above, enzymes in organisms have a variety of existing and potential uses in the practice of green chemistry. In many cases it is advantageous to isolate the enzyme used for a particular process from cells and use it outside the cellular environment. In a batch synthesis this can be done by mixing the enzyme with the reactants and allow it to catalyze the desired reaction making sure that optimum conditions of temperature and pH are maintained. This approach has several disadvantages. Enzymes are expensive and isolation of the enzyme from the reaction mixture is usually very costly and often not possible. Enzyme contamination of the product can cause difficulties.

The solution to the problem outlined above is often to employ **enzymeimmobilization** that uses a two-phase system in which the enzyme is in one phase and the reaction occurs in another. Sequestration of the enzyme in a separate phase enables its re-use or continuous use in flow-through systems and prevents enzyme contamination of the products. Several major techniques have been employed for enzyme immobilization including (1) adsorption onto a solid, (2) covalent binding onto a separate phase, (3) entrapment in a separate phase, (4) confinement with a membrane that allows transport of reactants and products but retains the enzyme. Ideally the matrix holding the enzyme should be capable of holding the enzyme as well as being inert, physically strong, chemically stable, and capable of being regenerated. The most common materials used to hold enzymes have been porous carbon, ion-exchange matrices, clays, polymeric resins, hydrous metal oxides, and glasses.

The procedure for immobilization of an enzyme begins with mixing the enzyme and the solid materials under suitable conditions of pH and ionic strength, sometimes along with binding agents. The support holding the immobilized enzyme is then incubated for some time. Finally, excess enzyme and, where used, binding agents are washed off of the support.

Rather than isolating enzymes and immobilizing them on a support, living microorganism cells, usually those of bacteria, are often used. Two main categories of immobilization are employed, attachment and entrapment. The simplest kind of immobilization is aggregation cross-linking in which the microbial cells form networks that compose their own support. This approach is usually confined to batch processes. Otherwise the major kinds of attachment immobilization are covalent binding, binding on ion-exchangers, adsorption binding, and biofilm formation. Probably the most common form of biofilm reactor is the trickling filter used for wastewater treatment (see Chapter 9) in which growing bacteria and protozoa form a film over solid support material (usually rock) and the wastewater is sprayed over the biofilm. This enables contact of the immobilized microorganisms with both the biodegradable material in the wastewater and atmospheric oxygen. Entrapment of microorganisms may be on organic polymer, on inorganic polymer, or behind a semi-permeable membrane.

Use of living organisms as sources of immobilized enzymes offers the advantage of not having to isolate the enzyme and, in cases where the organism is reproducing, of continuously replenishing the enzyme. A disadvantage includes having to maintain conditions under which the organism is viable. Also, living cells harbor numerous enzymes so side-reactions and unwanted products can be a problem.

Effects of Toxic Substances on Enzymes

Toxic substances may destroy enzymes or alter them so that they function improperly or not at all. Among the many toxic substances that act adversely with enzymes are heavy metals, cyanide, and various organic compounds such as insecticidal parathion. Many enzyme active sites through which an enzyme recognizes and bonds with a substrate contain -SH groups. Toxic heavy metal ions such as Pb^{2+} or Hg^{2+} are “sulfur seekers” that bind to the sulfur in the enzyme active site causing the enzyme to not function. A particularly potent class of toxic substances consists of the organophosphate “nerve gases” such as Sarin that inhibit the acetylcholinesterase enzyme required to stop nerve impulses. Very small doses of Sarin stop respiration by binding with acetylcholinesterase and causing it to not work. Discussed further in Section 7.9 under the topic of toxicological chemistry, toxicity to enzymes is a major consideration in the practice of green chemistry.

This page titled 2.5.7.7: Enzymes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **7.7: Enzymes** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.7.8: Biochemical Processes in Metabolism

So far, this chapter has discussed the cells in which biochemical processes occur, the major categories of biochemicals, and the enzymes that catalyze biochemical reactions. Biochemical processes involve the alteration of biomolecules, their synthesis, and their breakdown to provide the raw materials for new biomolecules, processes that fall under the category of **metabolism**. Metabolic processes may be divided into the two major categories of **anabolism** (synthesis) and **catabolism** (degradation of substances). An organism may use metabolic processes to yield energy or to modify the constituents of biomolecules. Metabolism is discussed here as it affects biochemicals and in Chapter 12, Section 12.4, as it applies to the function of organisms in the biosphere.

Metabolism is a very important consideration in green chemistry and sustainability. Toxic substances that impair metabolism pose a danger to humans and other organisms and attempts are made to avoid such substances in the practice of green chemistry. Exposures to environmental pollutants that impair metabolism endanger humans and other organisms; the control of such pollutants is an important aspect of environmental chemistry. Metabolic processes are used to make renewable raw materials and to modify substances to give desired materials. The complex metabolic process of photosynthesis provides the food that forms the base of essentially all food webs and is increasingly being called upon to provide renewable raw materials for manufacturing.

Energy-Yielding Processes

The processing of energy is obviously one of the most important metabolic functions of organisms. The metabolic processes by which organisms acquire and utilize energy are complex generally involving numerous steps and various enzymes. Organisms can process and utilize energy by one of three major processes:

- **Respiration** in which organic compounds undergo catabolism
- **Fermentation**, which differs from respiration in not having an electron transport chain
- **Photosynthesis**, in which light energy captured by plant and algal chloroplasts is used to synthesize sugars from carbon dioxide and water

There are two major pathways in respiration. **Oxic respiration**(called aerobic respiration in the older literature) requires molecular oxygen whereas **anoxic respiration** (anaerobic respiration) occurs in the absence of molecular oxygen. Oxic respiration uses the **Krebscycle** to obtain energy from the following reaction:

$$C_6H_{12}O_6 + 6\,O_2 \rightarrow 6\,CO_2 + 6\,H_2O + \text{energy}$$

About half of the energy released is converted to short-term stored chemical energy, particularly through the synthesis of **adenosine triphosphate**(**ATP**) shown in Figure 7.10

The highly energized ATP molecule is sometimes described as the “molecular unit of currency” for the transfer of energy within cells during metabolism. It releases its energy when it loses a phosphate group and reverts to adenosine diphosphate and other precursors. ATP is used by enzymes and proteins for cell processes including biosynthetic reactions (anabolism), cell division, and motility (such as occurs in moving protozoa cells). In so doing, ATP is continually being produced and reconverted back to its precursor species in an organism. Some studies have suggested that the human body processes its own mass in ATP during a single day! Whereas ATP is used for very short term energy storage and processing, for longer-term energy storage,glycogen or starch polysaccharides are synthesized, and for still longer-term energy storage, lipids(fats) are generated and retained by the organism.

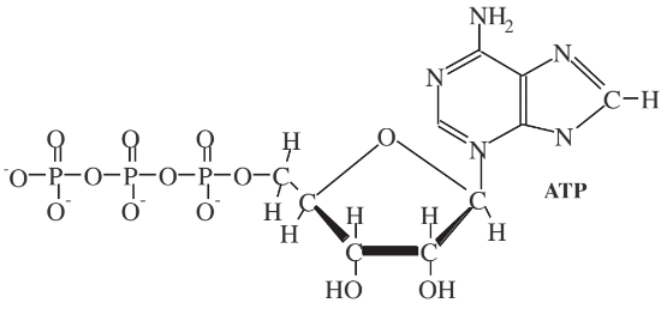


Figure 7.10. Adenosine triphosphate, a molecule strongly involved with energy transfer in living organisms.

As noted above, fermentation differs from respiration in not having an electron transport chain and organic compounds are the final electron acceptors rather than O2in the energy-yielding process. Many biochemical processes including some used to make commercial products are fermentations. A common example of fermentation is the production of ethanol from sugars by yeasts growing in the absence of molecular oxygen:

$$C_6H_{12}O_6 \rightarrow 2\,CO_2 + 2\,C_2H_5OH$$

Photosynthesis, is an energy-capture process in which light energy captured by plant and algal chloroplasts is used to synthesize sugars from carbon dioxide and water:

$$6\,CO_2 + 6\,H_2O + h\,\nu \rightarrow C_6H_{12}O_6 + 6\,O_2$$

When it is dark, plants cannot get the energy that they need from sunlight but still must carry on basic metabolic processes using stored food. Plant cells, like animal cells, contain mitochondria in which stored food is converted to energy by cellular respiration.

Nonphotosynthetic organisms depend upon organic matter produced by plants for their food and are said to be heterotrophic. They act as “middlemen” in the chemical reaction between oxygen and food material using the energy from the reaction to carry out their life processes Plant cells, which use sunlight as a source of energy and CO_2 as a source of carbon, are classified as **autotrophic**. In contrast, animal cells must depend upon organic material manufactured by plants for their food. These are called **heterotrophic** cells.

Biochemical conversions involving energy are very important in the practice of green chemistry and sustainability. The most obvious connection is the capture of solar energy as chemical energy by photosynthesis. As discussed in Chapter 15, photosynthetically produced biomass can serve as a source of chemically fixed carbon for the synthesis of chemical fuels including synthetic natural gas, gasoline, diesel fuel, and ethanol. A tantalizing possibility is to use recombinant DNA techniques to increase by several-fold the very low efficiency of photosynthesis by most plants. Fermentation has a strong role to play in sustainable energy development. As shown in Reaction 7.8.2, fermentation of glucose produces ethanol, which can be used as fuel.

Anoxic fermentation of biomass (abbreviated $\{CH_2O\}$) from sources such as sewage sludge or food wastes yields methane (natural gas) the cleanest-burning of all hydrocarbon fuels:

$$2\,(CH_2O) \rightarrow CH_4 + CO_2$$

This page titled 2.5.7.8: Biochemical Processes in Metabolism is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **7.8: Biochemical Processes in Metabolism** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.7.9: Biochemistry of Toxic Substances and Toxicological Chemistry

One of the more important aspects of biochemistry that is important in green chemistry is the way that organisms deal biochemically with toxic substances. There are two major aspects of this. One is the biochemical changes in toxic substances in an organism's system including the detoxification of such substances and in some cases the conversion of nontoxic compounds to toxic compounds. The second is the biochemical mode of action of toxic substances through which they exert a toxic effect. **Toxicological chemistry** is the science that deals with the chemical nature and reactions of toxic substances, including their origins, uses, and chemical aspects of exposure, fates, and disposal.[2] Toxicological chemistry addresses the relationships between the chemical properties and molecular structures of molecules and their toxicological effects. Figure 7.11 illustrates the definition of toxicological chemistry. This section discusses the biochemical aspects of toxicological chemistry. Toxicology, itself, and other details regarding toxicological chemistry are covered in Chapter 17.

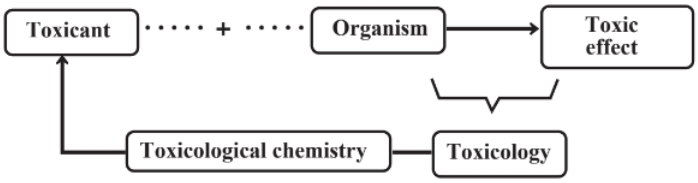


Figure 7.11. Toxicological chemistry is the biochemical science of toxic substances and their precursors. Toxicological chemistry relates the nature and reactions of these substances to toxicology

Biochemistry of Toxicants and **Protoxicants**

In some cases a toxic substance that enters into the system of a living organism is unchanged until it reacts to cause a toxic effect. This is the case with carbon monoxide, CO, which enters the bloodstream through the lungs and binds with blood hemoglobin to prevent oxygen transfer to tissues. In other case toxicants or their metabolic precursors (**protoxicants**) react in ways that may make them more toxic or that detoxify them and facilitate their elimination from the organism. **Xenobiotic compounds** are those that are normally foreign to living organisms. Some of the more toxic substances, such as the toxin produced by *Botulinus* bacteria or the venom of the deadly Australian inland taipan viper, are among the most toxic substances known. "Toxicant" is used here as a term to refer to toxic substances and their precursors including both xenobiotic materials and those of natural organisms. "The body" is used to refer to the human body, but also applies to other organisms as well.

Of particular importance in the metabolism of toxicants is **intermediaryxenobiotic metabolism** which results in the formation of somewhat transient species that are different from both those ingested and the ultimate product that is excreted. Intermediary metabolites may have significant toxicological effects. Toxicants and protoxicants in general are acted upon by enzymes that normally act upon an **endogenous substrate**, a material that is in the body naturally. For example, flavin-containing monooxygenase enzyme acts upon endogenous cysteamine to convert it to cystamine, but also functions to oxidize xenobiotic nitrogen and sulfur compounds.

Toxicants undergo **biotransformation** as a result of enzyme action, usually Phase I and Phase II reactions defined below. Some nonenzymatic transformations are also important including bonding of compounds with endogenous biochemical species without an enzyme catalyst, hydrolysis in body fluid media, or oxidation/reduction processes.

The likelihood of enzymatic metabolism in the body depends upon the physical and chemical properties of the species. Highly polar compounds, such as those that form ions readily, are less likely to enter the body system and generally are quickly excreted. Therefore, such compounds are unavailable, or only available for a short time, for enzymatic metabolism. Volatile compounds, such as dichloromethane or diethyl ether, are expelled so quickly from the lungs that enzymatic metabolism is less likely. This leaves as the most likely candidates for enzymatic metabolic reactions **nonpolar lipophilic compounds**, those that are relatively less soluble in aqueous biological fluids and more attracted to lipid species. Of these, the ones that are resistant to enzymatic attack (PCBs, for example) tend to bioaccumulate in lipid tissue.

Xenobiotic species may be metabolized in many body tissues and organs. The liver is of particular significance because materials entering systemic circulation from the gastrointestinal tract must first traverse the liver. As part of the body's defense against the entry of xenobiotic species, the most prominent sites of xenobiotic metabolism are those associated with entry into the body, such

as the skin and lungs. The gut wall through which xenobiotic species enter the body from the gastrointestinal tract is also a site of significant xenobiotic compound metabolism.

Phase I and Phase II Reactions

The metabolism of toxic substances may be divided into two phases. **Phase I reactions** normally consist of attachment of a functional group, usually accompanied by oxidation. For example, benzene, C_6H_6, (see Chapter 6, Section 6.2) is oxidized in the body by the action of the cytochrome P-450 enzyme system as shown in Figure 7.12.

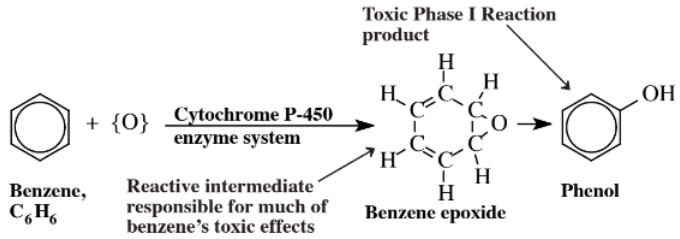


Figure 7.12. Phase I reaction of benzene to produce reactive intermediate benzene epoxide which rearranges to produce toxic phenol

The Phase I oxidation product of benzene is phenol, a toxic substance. A reactive intermediate in the process is benzene epoxide, which interacts with biomolecules to cause toxic effects. The phenol Phase I oxidation product of benzene may undergo a second reaction, a Phase II reaction in which it is bound with a conjugating agent that is endogenous to (produced naturally by) the body, such as glucuronide (Figure 7.13).

Although Phase I and Phase II reactions generally act to make xenobiotic substances more water soluble, more readily eliminated from the body, and less toxic, in some cases, the opposite occurs and metabolic processes make substances more toxic. Most known human carcinogens(cancer-causing agents) are actually produced by biochemical processes in the body from non-carcinogenic precursor substances.

Biochemical and Toxic Effects of Toxicants

Toxic substances, which, as noted above, are often produced by metabolic processes from nontoxic precursors, produce a toxic response by acting upon a **receptor** in the body. Typically, a receptor is an enzyme that is essential for some function in the body. As a consequence of the binding of the receptor to the toxicant there is a **biochemical effect**. A common example of a biochemical effect occurs when a toxicant binds to an enzyme such that the bound enzyme may be inhibited from carrying out its normal function. As a result of a biochemical effect, there is a response, such as a behavioral or physiological response, which constitutes the actual observed toxic effect. Acetylcholinesterase enzyme inhibited by binding to nerve gas Sarin may fail to stop nerve impulses in breathing processes, leading to asphyxiation. The phenomena just described occur in the **dynamicphase** of toxicant action as summarized in Figure 7.14.

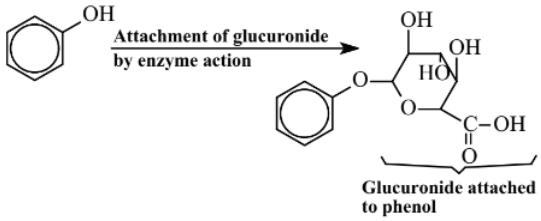


Figure 7.13. A Phase II reaction in which phenol is conjugated with glucuronide, an endogenous conjugating agent.

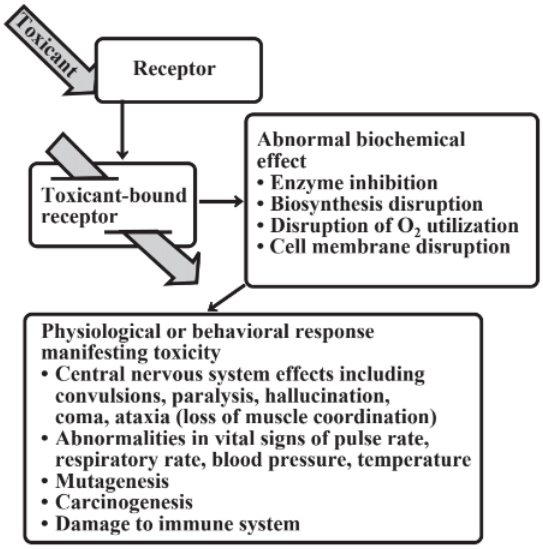


Figure 7.14. The dynamic phase of toxicant action in which a toxicant interacts with an endogenous receptor leading to biochemical changes that adversely affect an organism

Toxicological Chemistry and the Endocrine System The endocrine glands and the hormones they produce (see Section 7.5 and Figure 7.6) are important in the consideration of toxicological chemistry, green chemistry and sustainability. Of particular importance are substances that organisms get exposed to through their environment, food, and drinking water that have the potential to disrupt the crucial endocrine gland activities that regulate the metabolism and reproductive functions of organisms. **Hormonally active agents** exhibit hormone-like activity that may be detrimental. Most commonly these are estrogenic substances that act like the female sex hormone estrogen. Some of these survive water treatment processes. Discharged to receiving waters they affect aquatic organisms and potentially can get into water that humans drink. Among such substances are estrogen, an endogenous sex hormone; 17a-ethinylestradiol, an ingredient of oral contraceptives; and chemicals from industrial and consumer sources that mimic estrogen. Estrogenic substances from artificial sources are called **xenoestrogens** and include antioxidants, bisphenol A, dioxins, PCBs, phytoestrogens (from plants), some pesticides (chlordecone, dieldrin, DDT and its metabolites, methoxychlor, toxaphene), preservatives, and phthalic acid esters (butylbenzyl phthalate).

The practice of green chemistry minimizes the production and use of xenoestrogens and attempts to prevent their introduction into the environment. A particular need exists for the development of alternatives to xenoestrogenic bisphenol A and phthalate plasticizers. These substances improve the properties of plastics and as molecules much smaller than those in the plastic polymer tend to get into the environment and food chain. Modification of the plastic polymer formulation to give desired properties without the need for plasticizer additives would be especially desirable.

This page titled 2.5.7.9: Biochemistry of Toxic Substances and Toxicological Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **7.9: Biochemistry of Toxic Substances and Toxicological Chemistry** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.7.10: Literature Cited

1. Voet, Donald, Judith G. Voet, and Charlotte W. Pratt, *Fundamentals of Biochemistry*, Wiley, New York, 2008.

2. Stanley E. Manahan, *Toxicological Chemistry and Biochemistry*, 3rd ed., Taylor & Francis/CRC Press, Boca Raton, Florida, 2002.

3. Rogers, Kara, Ed., *The Cell*, Rosen Educational Services, New York, 2011.

2.5.7.10: Literature Cited is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Literature Cited** is licensed CC BY-NC-SA 4.0.

2.5.7.11: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. What are the basic building blocks of proteins, and how do they determine the primary structure of proteins?

2. What is meant by denaturation of proteins? Is it bad?

3. What are some major kinds of proteins?

4. What is the approximate simple formula of carbohydrates?

5. Fill in the blanks of the following pertaining to carbohydrates: Glucose is an example of a_____, sucrose is a _____, and starch and cellulose areboth _____.

6. How are lipids defined and how does this definition differ from that of other biomolecules?

7. What does DNA stand for? What are 6 specific ingredients of DNA?

8. Although lipids are defined by a physical property that they all share, what is a common characteristic of lipid structure?

9. From the structures given in Figure 7.5, what kind of functional group seems to be common in lipids?

10. How does the compatibility of lipids with organic substances, such as organochlorine compounds, influence the environmental behavior of such compounds?

11. What distinguishes RNA from DNA? How are they similar?

12. What are the three constituents of all basic units of nucleic acids?

13. Exposure of a person to toxic benzene can be estimated by measuring phenol in blood. Explain the rationale for such an analysis. Why is benzene epoxide not commonly determined to estimate benzene exposure?

14. Consider the toxicity of inhaled carbon monoxide in the context of Figure 7.14. Identify for carbon monoxide the receptor, the abnormal biochemical effect, and the physiological response manifesting toxicity.

15. What is the toxicological importance of lipids? How are lipids related to hydrophobic pollutants and toxicants?

16. What is the function of a hydrolase enzyme?

17. Match the cell structure on the left with its function on the right, below:

    A. Mitochondria 1. Toxicant metabolism

    B. Endoplasmic reticulum 2. Fills the cell

    C. Cell membrane 3. Deoxyribonucleic acid

    D. Cytoplasm 4. Mediate energy conversion and utilization

    E. Cell nucleus 5. Encloses the cell and regulates the passage of materials into and out of the cell interior

18. The formula of simple sugars is $C_6H_{12}O_6$. The simple formula of higher carbohydrates is $C_6H_{10}O_5$. Of course, many of these units are required to make a molecule of starch or cellulose. If higher carbohydrates are formed by joining together molecules of simple sugars, why is there a difference in the ratios of C, H, and O atoms in the higher carbohydrates as compared to the simple sugars? What do these formulas suggest about the kind of enzyme that would be required to produce glucose from a higher carbohydrate such as cellulose?

19. What would be the chemical formula of a *tri*saccharide made by the bonding together of three simple sugar molecules?

20. The general formula of cellulose may be represented as $(C_6H_{10}O_5)_x$. If the molar mass of a molecule of cellulose is 400,000, what is the estimated value of$x$?

21. Glycine and phenylalanine can join together to form two different dipeptides. What are the structures of these two dipeptides?

22. Fungi, which break down wood, straw, and other plant material, have what are called“exoenzymes.” Fungi have no teeth and cannot break up plant material physically by force. Knowing this, what do you suppose an exoenzyme is? Explain how you think it might operate in the process by which fungi break down something as tough as wood.

23. The straight-chain alcohol with 10 carbons is called decanol. What do you think would be the formula of decyl stearate? To what class of compounds would it belong?

24. In what respect is an enzyme and its substrate like two opposite strands of DNA?

2.5.7.11: Questions and Problems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** is licensed CC BY-NC-SA 4.0.

2.5.7.12: Supplementary References

Tymoczko, John L., Jeremy Mark Berg, and Lubert Stryer, *Biochemistry: A Short Course*, 6th ed.,W.H. Freeman, New York, 2009.

Bettelheim, Frederick A., William H. Brown, and Jerry March, *Introduction to Organic and Biochemistry*, 7th ed., Brooks/Cole, Belmont, CA, 2009.

Chesworth, J. M., T. Stuchbury, and J.R. Scaife, *An Introduction to Agricultural Biochemistry*, Chapman and Hall, London, 1998.

Elliott, William H., and Daphne C. Elliott, *Biochemistry and Molecular Biology*, 4th ed., Oxford University Press, New York, 2009.

Garrett, Reginald H., Charles M. Grisham, *Biochemistry*, 4th ed., Brooks/Cole, Belmont, CA,2009.

Kuchel, Philip W., Simon Easterbrook-Smith, Vanessa Gysbers, J. Mitchell Guss, Dale P.Hancock, Jill M. Johnston, Alan Jones, and Jacqui M. Matthews, *Schaum's Outline of Biochemistry*, McGraw-Hill, Boston, 2010.

Bowsher, Caroline, Martin Steer, and Alyson Tobin *Plant Biochemistry*, Garland Science, New York, 2008.

Markandey, D. K., and N. Rajvaida, *Environmental Biochemistry*, APH Publishing, New Dehli,2005.

McKee, Trudy, and James R. McKee, *Biochemistry: The Molecular Basis of Life*, 4th ed., Oxford University Press, New York, 2009.

Moore, John, T., and Richard Langley, *Biochemistry for Dummies*, Wiley Publishing, Inc., Indianapolis, IN, 2008.

Pratt, Charlotte W., and Kathleen Cornely, *Essential Biochemistry*, Wiley, Hoboken, NJ, 2011.

Swanson, Todd A., Sandra I. Kim, and Marc J. Glucksman, *Biochemistry and Molecular Biology*, 4th ed., Lippincott Williams & Wilkins, Philadelphia, 2007.

Tymoczko, John L., Jeremy M. Berg, and Lubert Stryer, *Biochemistry: A Short Course*, W.H.Freeman and Co., New York, 2009.

Voet, Donald, Judith G. Voet, *Fundamentals of Biochemistry: Life at the Molecular Level*, 3rd ed., Wiley, Hoboken, NJ, 2008.

2.5.7.12: Supplementary References is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Supplementary References** is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

2.5.8: The Five Environmental Spheres and Biogeochemical Cycle

“In addition to water, air, solid earth, and life, a fifth sphere of the environment must be considered, the anthrosphere consisting of the things made and used by humans. The anthrosphere has such a profound influence that we are now entering a new epoch, the anthropocene in which Earth’s environment will be largely determined by human activities.

2.5.8.1: Five Environmental Spheres

2.5.8.2: The Hydrosphere

2.5.8.3: The Atmosphere

2.5.8.4: The Geosphere

2.5.8.5: The Biosphere

2.5.8.6: The Anthrosphere

2.5.8.7: Cycles of Matter

2.5.8.8: Questions and Problems

2.5.8.9: Supplementary References

This page titled 2.5.8: The Five Environmental Spheres and Biogeochemical Cycle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.8.1: Five Environmental Spheres

In Section 1.2 it was noted that Earth's environment may be regarded as consisting of five spheres: (1) The **hydrosphere**, (2) the **atmosphere**, (3) the **geosphere**, (4) the **biosphere**, and (5) the **anthrosphere** (that part of the environment constructed and operated by humans). All of these spheres are closely interconnected and interactive, continuously exchanging matter and energy and influencing each other. Much of this interaction is through biogeochemical cycles in which biological, geochemical, aquatic, atmospheric, and, increasingly, anthrospheric processes act to exchange matter and energy among the five environmental spheres and determine Earth'senvironment. The next several chapters address these five environmental spheres and the cycles in which they are involved.

Consideration of the five environmental spheres and their interactions through biogeochemical cycles are very important in green chemistry, science, and technology as well as sustainability as a whole. Earth's natural capital (see Section 1.4) resides in the four "natural" spheres and is utilized (sometimes shamelessly exploited) and managed in the anthrosphere. The interfaces between spheres are particularly important for, among other reasons, they are where materials and energy are exchanged. As discussed in Section 1.2, one such interface is the one between the geosphere and the atmosphere where most plants grow that support life on Earth. Like most such interfaces, it is very thin compared to the spheres themselves, generally extending into the geosphere for only the meter or less penetrated by plant roots and into the atmosphere only to the height of the plants. The boundary between the anthrosphere and the atmosphere is where air pollutants from automobile engines and other sources enter the atmosphere. Many other examples of important interfaces could be given.

This page titled 2.5.8.1: Five Environmental Spheres is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **8.1: Five Environmental Spheres** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.8.2: The Hydrosphere

Water, chemical formula H_2O, comprises the hydrosphere (Figure 8.1). As discussed in more detail in Chapter 9, although it has a simple chemical formula, water is actually a very complex substance largely because of the water molecule's polarity and ability to form hydrogen bonds. Water participates in one of the great natural cycles of matter, the **hydrologiccycle** illustrated in Figure 8.1. Basically the hydrologic cycle is powered by solar energy that evaporates water as atmospheric water vapor from the oceans and bodies of fresh water from where it may be

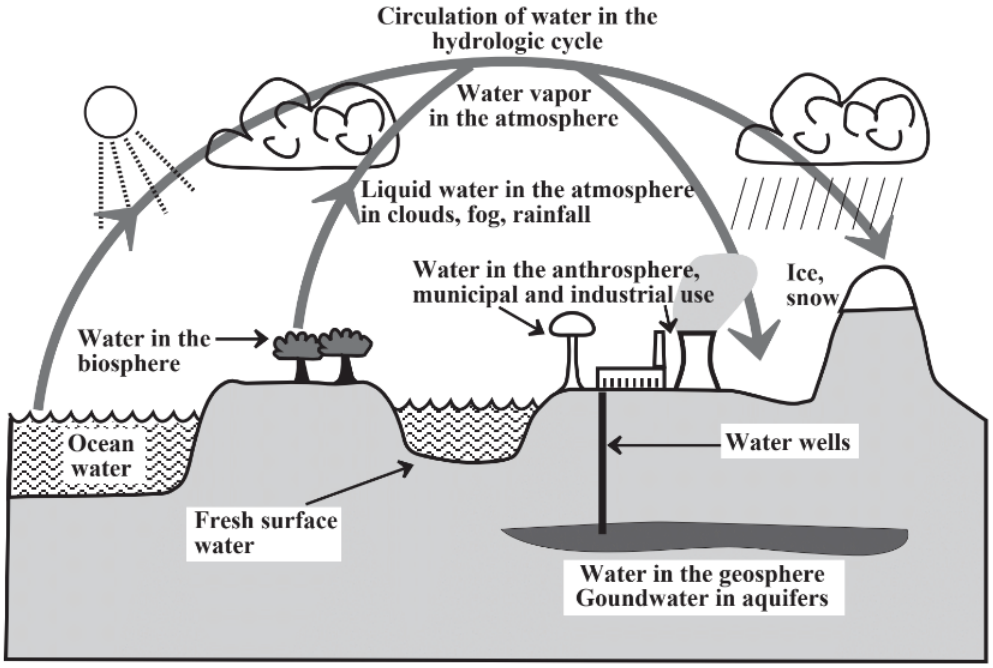


Figure 8.1. The hydrosphere overlaps strongly with all the other environmental spheres. This illustration shows water in bodies of water, underground as groundwater, in snow pack, in plants in the biosphere, as droplets and vapor in the atmosphere, and in water distribution systems and cooling towers in the anthrosphere. Water's cycle in the environment is the hydrologic cycle shown here in which solar energy evaporates water from the oceans, surface bodies of water, soil, and plants (transpiration) and the water in the atmosphere is carried for a distance (sometimes thousands of kilometers), falls back to Earth as precipitation, infiltrates in part into groundwater, moves on Earth's surface in rivers, back in part to the oceans, then evaporates to renew the cycle.

carried by wind currents through the atmosphere to fall as rain, snow, or other forms of precipitation in areas far from the source. In addition to carrying water, the hydrologic cycle conveys energy absorbed as latent heat when water is evaporated by solar energy and released as heat when the water condenses to form precipitation.

The hydrosphere is shown in Figure 8.1, which also illustrates its relationship to the other environmental spheres. A remarkable 97.5% of Earth's water is saltwater in the oceans. Of the remaining fresh water, 1.7% of Earth's total water is in polar ice and the Greenland icecap. This leaves only about 0.77% of Earth's water as fresh water for other than ocean-dwelling organisms and for use by humans. This fresh water occurs in natural lakes, rivers, impoundments made by humans, and underground as groundwater. At any given time a miniscule, but very important fraction is contained in the anthrosphere, such as in water distribution systems.

Water is the most widely used substance by organisms in the anthrosphere. Humans use water in their households for drinking, food preparation, cleaning, and disposal of wastes, drawing water from rivers, from impoundments made by damming rivers, and by pumping from underground aquifers. Moving water is one of the oldest forms of power harnessed by humans. Water wheels date back more than 2000 years and hydroelectric power is still the leading source of renewable energy. Hot water vapor (steam) is widely used for heat transfer in industry and in buildings and is the largest means of electrical power generation through the mechanism of steam turbines coupled to electrical generators.

Sustainability demands consideration of the water resource, shortages of which resulting from climate-induced droughts have caused severe problems for many organisms and declines of major civilizations. Devastating floods displace and even kill large

numbers of people throughout the world and destroy homes and other structures. Severe droughts curtail plant productivity resulting in food shortages for humans and animals in natural ecosystems and often necessitating slaughter of farm animals. It is feared that both drought and the severity of occasional flooding will become much worse as the result of global warming brought on by rising carbon dioxide levels in the atmosphere (see Chapter 10, Section 10.6).

The maintenance of healthy and prosperous human populations requires consideration of both water quality and quantity. Waterborne diseases including cholera and typhoid have killed millions of people and these and others, especially dysentery, are problems in less developed areas lacking proper sanitation. The prevention of water pollution has been a major objective of the environmental movement and avoiding discharge of harmful water pollutant chemicals is one of the main objectives of the practice of green technology. Water supplies are a concern with respect to terrorism because of their potential for deliberate contamination with biological or chemical agents.

Since ancient times humans have built water reservoirs for water storage and dikes and dams for flood control. Such water management measures have enabled development of large populations in arid regions and in areas vulnerable to flooding. However, unusually severe, prolonged droughts do occur and in past times entire civilizations have been wiped out as a result. The effects of severe droughts are exacerbated by the fact that control of water supplies has enabled excessive growth in water-deficient areas. The Las Vegas metropolitan region of the U.S.and Mexico City in Mexico are examples of metropolitan regions that have outgrown the natural water capital available to them.

Floods cause hundreds of millions of dollars in damage to communities where construction of river dikes and impoundments have enabled agricultural and other developments in flood-prone areas that become overwhelmed by “hundred-year” flood events. Such an incident took place along the Missouri River in 1993 when a 500-year flood overwhelmed most of the protective structures. Failure of the protective systems caused much greater devastation than would otherwise have been the case when Hurricane Katrina destroyed much of New Orleans in 2005. Following the 1993 Missouri River flood, sensible actions were taken in some areas where farm property along the Missouri river was purchased by government agencies and allowed to revert to wildlife habitat in its natural state, which included periodic flooding. It would have been sensible for the future of New Orleans to move areas located below sea level and flooded by levee failure in 2005 to higher ground and to avoid trying to thwart the natural tendency of water to seek lower levels where humans may try to live.

Problems with water supply are discussed in Chapter 9. Figure 9.2 showing rainfall patterns in the continental U.S. reveals that the eastern continental U.S. receives generally adequate rainfall, although damaging droughts may still occur in that region. However, except for northern coastal regions, the western half of the continental U.S. is water-deficient. Water-deficient areas of theU.S. including southern California, Arizona, Nevada, and Colorado have exhibited some of the most rapid population growth in recent decades putting increasing pressure on limited water supplies and making the region vulnerable to prolonged severe droughts. Even much more severe water supply problems exist in other parts of the world, such as sections of Africa and the Middle East including the area of Palestine and Israel.

Damage to the Hydrosphere

Earth could have lost most of its water by now except for one very fortunate atmospheric feature, the very cold tropopause boundary at the upper part of the atmospheric troposphere. At a temperature well below the freezing point of water, this region converts water vapor to ice that remains in the troposphere and participates in the hydrologic cycle. Were this not the case, the water vapor would infiltrate the next higher atmospheric layer, the stratosphere, where highly energetic solar ultraviolet radiation would split H atoms off the H_2O molecules. These very light atoms and H_2 molecules formed from them would have diffused into space, leaving Earth with an arid Martian landscape. In fact, there is probably a net influx of water into Earth’s atmosphere from meteorites that are largely composed of water.

Although water is not destroyed on Earth, the hydrosphere certainly can suffer damage by human activities. One of the main ones of these is excessive utilization of water in arid regions. Withdrawal of irrigation water from rivers in arid regions has reduced some once mighty rivers to trickles by the time they reach the ocean. The water is not destroyed, but it evaporates and in some cases infiltrates to below ground.

Mexico City manifests one of the most serious problems of water over-use, the depletion of groundwater. This highly populous city was built on an old lake bed, and excessive pumping of ground water has caused land subsidence and damage to surface structures.

In the U.S. wasteful use of groundwater is illustrated by the depletion of the High Plains Aquifer (Figure 8.2) commonly called the Ogallala aquifer. Largely composed of fossil water remaining from the last

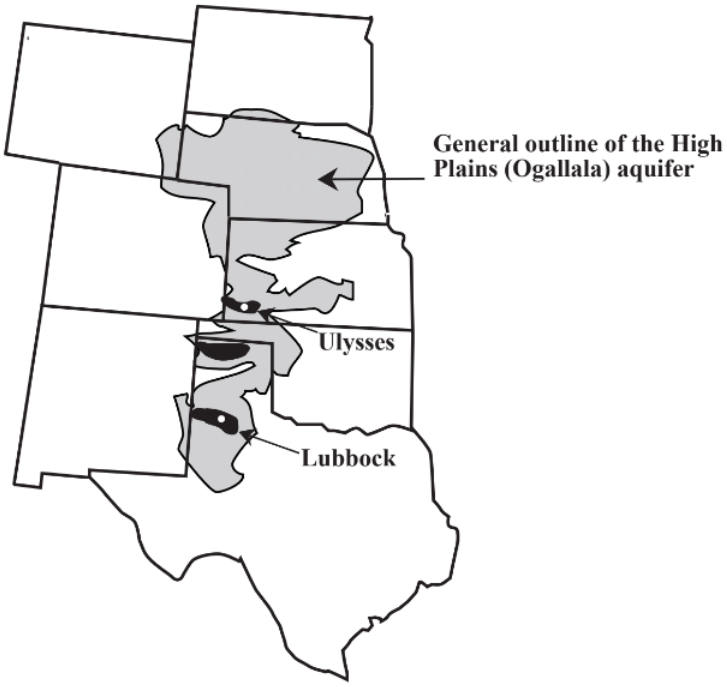


Figure 8.2. The water of the vast Ogallala aquifer, a precious resource remaining from the last Ice Age, has been recklessly exploited and depleted to grow water-intensive crops in semi-arid regions that do not normally support such crops. Areas outlined in black are those in southwestern Kansas and the Texas Panhandle where depletion has been especially severe.

Ice Age, the Ogallala aquifer lies beneath much of Nebraska, western Kansas, the Oklahoma and Texas panhandles, and small sections of eastern Wyoming, Colorado, and New Mexico. Although it is recharged from surface water in parts of Nebraska, it is largely composed of fossil water from the last Ice Age. It contains an astounding amount of water, enough to cover the entire United States to a depth of around 1.5 meters!

Since the 1940s, huge quantities of water pumped from the Ogallala aquifer have been used to irrigate corn and other crops not normally adapted to the High Plains region. As a result, the water table (level reached by water in a well drilled into an aquifer, Figure 8.3) has dropped dramatically, exceeding 50 meters in some areas, In the single decade beginning in 1995, the water level dropped by 6 meters in the middle of Kansas' irrigated corn belt around Ulysses, Kansas. Such unsustainable water depletion will force a shift from thirsty crops, such as corn, to those that require less water, such as milo.

https://bio.libretexts.org/@go/page/108834

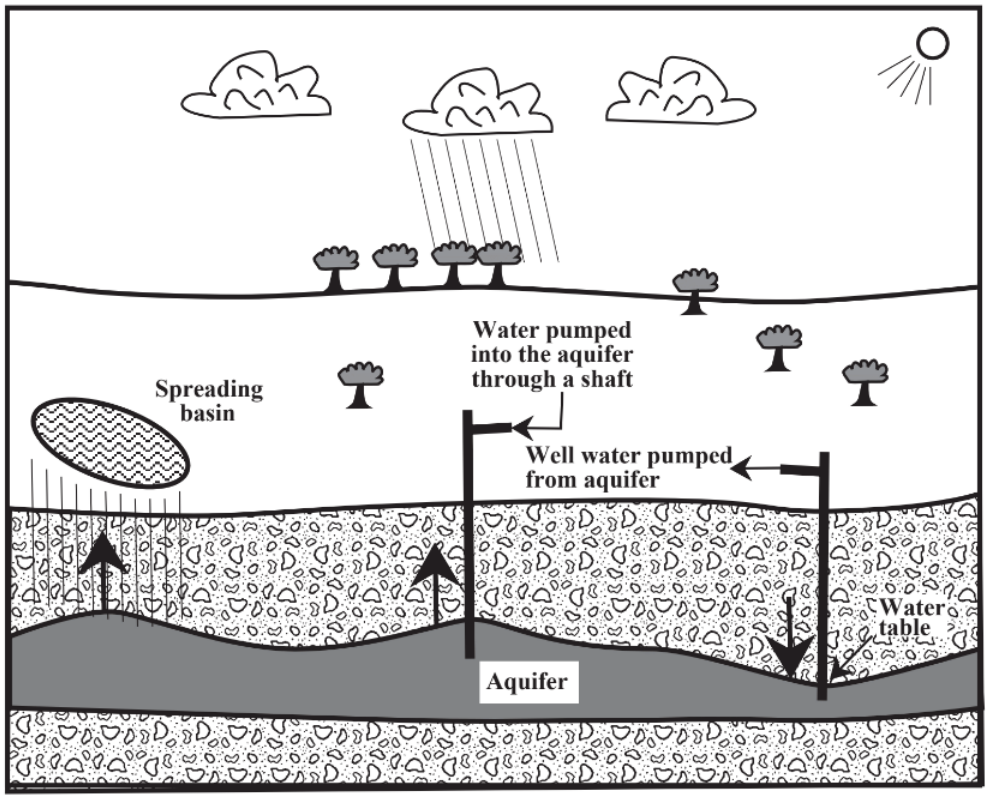


Figure 8.3. Water levels in an aquifer can be directly influenced by human activities. Water levels are lowered by withdrawal of water from wells. The recharge of aquifers can be aided by water infiltrating from spreading basins on the surface or by direct injection of water

The depletion of water supplies has numerous implications for sustainability. Clearly, the U.S. can meet its food demands without exploiting the Ogalla aquifer and other diminishing sources of water. A major thrust of the U.S. energy plan — such as it is — has been increased production of biofuels — ethanol from fermentation of glucose sugar derived from corn and biodiesel fuel from soybeans. The voracious consumption of irrigation water required to grow enough of these crops to make a significant difference in fuel resources is not sustainable. Some underground water supplies should be regarded as depletable resources and reserved for municipal water supplies and manufacturing, which require only a fraction as much water as does irrigation.

Sustaining the Hydrosphere

Fortunately, the solar-powered hydrologic cycle and the inexhaustible supply of water in Earth's oceans makes water one of nature's most renewable resources. Water is never really consumed or destroyed, although it may become so polluted or dispersed that its reclamation is impractical. Even water infiltrating into the ground may be regarded as recycled water because it renews groundwater sources. Although the thought of so doing has largely prevented efforts to completely recycle water that has been through domestic sewage systems, recycling of this water following purification will have to be practiced in some water-deficient areas in the future. In fact, recycling sewage water has long been the practice where municipalities take their water supplies from rivers into which other municipalities discharge treated wastewater. A very favorable development during the last 30 years is that the trend toward ever increasing water use that prevailed in the U.S. until about 1980 has leveled off due to more efficient irrigation and industrial processes and it appears that total utilization of water will remain relatively constant in the U.S into the future.

As exemplified by the exploitation of the Ogallala aquifer discussed above, human manipulation of the hydrosphere to provide water has often had adverse effects. However, application of the principles of green science and technology can enable supply of water to water-deficient areas without damaging the environment and even with enhancement of water quality. Unlike the over-utilized Colorado and Rio Grande rivers of the southwestern U.S., the enormous flow of the Mississippi River could be tapped for water-deficient areas. One scheme would be to divert a fraction of this flow near the mouth of the Mississippi where it discharges to the Gulf of Mexico and pump the water using abundant wind power to arid regions of the U.S. Southwest and northern Mexico. Mississippi River water retained for some time in constructed wetlands near the point of the diversion could undergo self-

purification, collecting sediment that builds up land mass and removing nutrients that are now harmful to water quality in the Gulf of Mexico. Aquatic plants growing in the wetlands and thriving on fertilizer runoff from the upstream Mississippi watershed could remove nutrients that are now harmful to water quality in the Gulf of Mexico.

Groundwater recharge is another key to water sustainability that is being practiced in parts of the world. Most groundwater recharge occurs naturally, although it has been reduced by paving and surface modifications in the anthrosphere. Anthrospheric constructs on the surface can be designed to maximize recharge. For example, some paving surfaces in China have been made of porous materials that allow water to penetrate into the ground below. Two of the more active approaches to groundwater recharge are shown in Figure 8.3. One of these is water pumped into a shaft that extends underground and even to the aquifer itself. Surfaces of these conduits can become clogged with silt, bacterial growths, and other materials suspended in the recharge water and may have to be cleaned periodically. A spreading basin consists of a reservoir of water excavated into porous geospheric material from which water flows into the aquifer. An advantage is the purification of water that occurs through contact with mineral matter, but the process does not work well if aquifers are overlain by poorly pervious layers

---

This page titled 2.5.8.2: The Hydrosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **8.2: The Hydrosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.8.3: The Atmosphere

Illustrated in Figure 8.4, the atmosphere having a total mass of about 5.15×10^{15} tons is a layer of gases blanketing Earth that diminishes rapidly in density with increasing altitude. More than 99% of the atmosphere's mass is within 40 kilometers (km) of Earth's surface, with the majority of the air below 10 km altitude (compared to Earth's diameter of almost 13,000 km). A person exposed to air at the approximately 13,000 m altitude at which commercial jet aircraft fly could remain conscious for only about 15 seconds without supplementary oxygen. There is no clearly defined upper limit to the atmosphere, which keeps getting thinner with increasing altitude. A practical upper limit may be considered to be an altitude of about 1000 km above which airmolecules can be lost to space (a region called the exosphere).

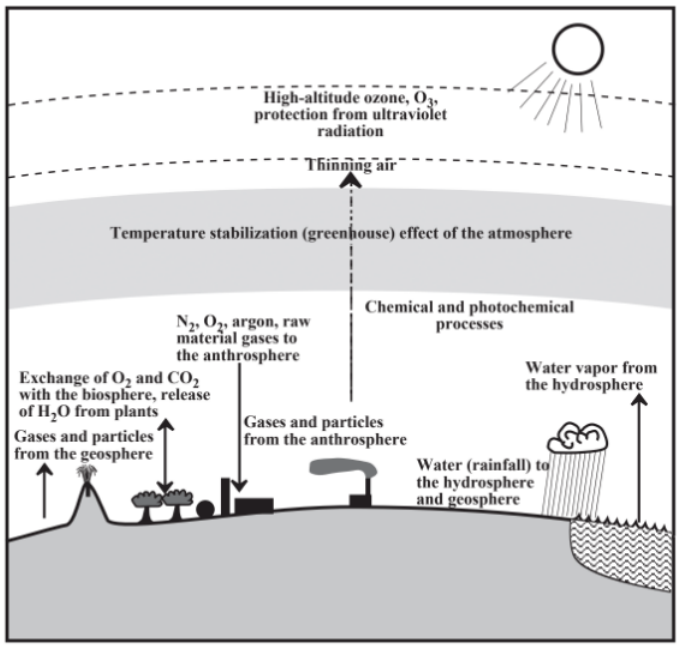


Figure 8.4. The atmosphere is a layer of gas, most of which is concentrated in a very thin region just above Earth's surface. It serves as a source of some essential materials, protects life on Earth from deadly electromagnetic radiation, transports water as part of the hydrologic cycle, participates in the carbon, oxygen, and sulfur cycles, and is an essential part of natural capital

The atmosphere nurtures life on Earth in many important respects. Some of the main ones of these are the following:

• The atmosphere constitutes much of Earth's natural capital because of its attributes listed below.

• The atmosphere is a source of molecular O_2 for all organisms that require it including humans and all other animals. In addition, pure oxygen, argon, and neon are extracted from the atmosphere for industrial uses.

• At approximately 0.039% carbon dioxide, CO_2, the atmosphere is the source of carbon that plants and other photosynthetic organisms use to synthesize biomass

• Consisting mostly of molecular N_2, the atmosphere serves as a source of nitrogen that is an essential component of protein and other biochemicals as well as a constituent of a variety of synthetic chemicals. Organisms "fix" this nitrogen in the biosphere chemically by the action of bacteria such as *Rhizobium* and it is fixed synthetically in the anthrosphere under much more severe conditions of temperature and pressure.

•The atmosphere acts as a blanket to keep Earth's surface at an average temperature of about 15° C at sea level and within a temperature range that enables life to exist(the "good" greenhouse effect).

•Earth's atmosphere absorbs very short wavelength ultraviolet radiation from the sun and space, which, if it reached organisms on Earth's surface, would tear apart the complex biomolecules essential for life. In this respect, the stratospheric ozone layer is of particular importance

• The atmosphere contains and carries water vapor evaporated from oceans that forms rain and other kinds of precipitation over land in the hydrologic cycle(Figure 8.1)

The atmosphere is 1–3% by volume water vapor, a level that is somewhat higher in tropical regions and lower in desert areas and at higher altitudes where condensation to liquid droplets and ice crystals removes water vapor. Exclusive of water vapor, on a dry basis air is 78.1% by volume nitrogen gas (N_2), 21.0% O_2, 0.9% noble gas argon, and almost 0.039% CO_2, a level that keeps increasing by a little more than 0.001% per year. In addition, there are numerous trace gases in the atmosphere at levels below 0.002% including ammonia, carbon monoxide, helium, hydrogen, krypton, methane, neon, nitrogen dioxide, nitrous oxide, ozone, sulfur dioxide, and xenon.

Figure 8.4 shows some of the main features and aspects of the atmosphere and its relationship to other environmental spheres. Except for a few aviators who fly briefly into the stratosphere, living organisms experience the lowest layer called the **troposphere** characterized by decreasing temperature and density with increasing altitude. The troposphere extends from surface level, where the average temperature is 15°C, to about 11 km (the approximate cruising altitude of commercial jet aircraft) where the average temperature is –56°C (the tropopause discussed in Section 8.2)

Above the troposphere is the **stratosphere** in which the average temperature increases from about –56°C at its lower boundary to about –2°C at its upper limit. The stratosphere is warmed by energy of intense solar radiation impinging on air molecules. Because this radiation can break the bonds holding O_2 molecules together, at higher altitudes the stratosphere maintains a significant level of O atoms and of ozone (O_3) molecules formed by combination of O atoms with O_2 molecules. Stratospheric ozone molecules are essential for the ability of humans and other organisms to exist on Earth's surface because of their ability to filter out damaging ultraviolet radiation before it can penetrate to Earth's surface. Although the stratosphere's ozone is spread over many km in altitude, it is commonly called the **ozone layer**. If all this ozone that is so essential for life were in a single layer of pure ozone at conditions near Earth's surface, it would be only about 3 millimeters thick! Some classes of chemical species, especially the chlorofluorocarbons or Freons formerly used as refrigerants, are known to react in ways that destroy stratospheric ozone, and their elimination from commerce has been one of the major objectives of efforts in achieving sustainability.

Above the stratosphere are the atmospheric mesosphere and thermosphere. which are relatively less important in the discussion of the atmosphere. Radiation energetic enough to tear electrons away from atmospheric molecules and atoms reaches these regions giving rise to a region containing ions called the **ionosphere.**

Earth's atmosphere is crucial in absorbing, distributing, and radiating the enormous amount of energy that comes from the sun. A square meter of surface directly exposed to sunlight unfiltered by air would receive energy from the sun at a power level of 1,340 watts. Called the **solar flux**, this level of power impinging on just one square meter could power an electric iron or thirteen 100-watt light bulbs plus a 40-watt bulb! If one considers Earth's cross-sectional area, the rate of total incoming solar energy is huge. Incoming solar radiation is in the form of electromagnetic radiation, which has a wavelike character in which shorter wavelengths are more energetic. The incoming radiation in the form electromagnetic radiation centered in the visible wavelength region with a maximum intensity at a wavelength of 500 nanometers (1 nm = 10^{-9}m) is largely absorbed and converted to heat in the atmosphere and at Earth's surface.

On average, the incoming solar energy must be balanced with heat energy radiated back into space; otherwise Earth would have melted and vaporized long ago. The outbound heat energy radiates into space as infrared radiation between about 2 micrometers and 40 μm (1 μm = 10^{-6} m)with a maximum intensity at about 10 μm). This energy is delayed on its outward path by being re-absorbed by water molecules, carbon dioxide, methane, and other minor species in the atmosphere. This has a warming (greenhouse) effect that is very important in sustaining life on Earth. As discussed in Chapter 10, anthrospheric discharges of greenhouse gases, especially carbon dioxide and methane, are likely causing an excessive greenhouse effect, which will have harmful effects on global climate.

Climate

Largely determined by conditions in the atmosphere, **climate** is crucial to the well-being of humans and other organisms on Earth. **Weather** refers to such factors as rain, wind, cloud cover, atmospheric pressure, and temperature whereas climate involves these conditions over a long period of time such as the warm, low-humidity weather that prevails in southern California or the cool,

generally rainy conditions of Ireland. **Meteorology** is the science of the atmosphere and weather and climate is addressed by **climatology.**

Much of the driving force behind weather and climate is due to the fact that the incoming flux of solar energy is very intense in regions around the equator and very low in polar regions due to the angles at which the solar flux impinges Earth in these regions. Heated equatorial air tends to expand and flow away from the equatorial regions, creating winds and carrying with it large quantities of energy and water vapor evaporated from the oceans. As the air cools, water vapor condenses forming precipitation and warming the air from heat released when water goes from a vapor to a liquid. This process is the driving force behind hurricanes and typhoons which can result in torrential rainfalls and damaging winds.

Meteorological phenomena have an important influence on air pollution. An important example occurs with temperature inversions in which a layer of relatively warm air confines a surface layer of somewhat cooler, more dense stagnant air close to the ground. Hydrocarbons and nitrogen oxides confined in the stagnant air mass are acted upon by solar energy to cook up the noxious mixture of ozone, oxidants, aldehydes and other unpleasant materials that constitute photochemical smog.

Winds are important in air pollution. The lack of substantial wind is required for the formation of photochemical smog. Sulfur dioxide emitted to the atmosphere may be relatively innocuous near the point of discharge but transformed to harmful acid rain as it is carried some distance by wind. That is the reason that New England and parts of Canada can be afflicted with acid rain from sulfur dioxide given off by coal-fired power plants some distance away in the Ohio River Valley.

Climate and sustainability go hand-in-hand. Favorable climate conditions and a relatively unpolluted atmosphere are important parts of natural capital. A favorable climate is required to maintain food productivity. One of the greatest concerns with the emission of excessive amounts of greenhouse gases to the atmosphere is warming that will result in catastrophic drought in formerly productive agricultural regions leading to much reduced food production and even widespread starvation.

Anthrospheric Influences on the Atmosphere

Human activities in the anthrosphere strongly influence the atmosphere and have enormous potential to affect Earth's environment as a whole. Photochemical smog with its destructive ozone and other oxidants along with other pollutants and visibility-obscuring atmospheric particulate matter results when nitrogen oxides and reactive hydrocarbons, largely from internal combustion vehicle engine exhausts, are emitted to the atmosphere from the anthrosphere. The sulfur dioxide and nitrogen oxide precursors to strong sulfuric and nitric acids in acid rain come from anthrospheric activities associated with fossil fuel combustion and other sources such as roasting of sulfide metal ores. Because of factors such as these, one of the most important aspects of sustainability is the construction and operation of the anthrosphere in ways that preserve the quality of the atmosphere. One of the important aspects of green chemistry is to use products and processes that do not contribute to damage to the atmosphere.

This page titled 2.5.8.3: The Atmosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **8.3: The Atmosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.8.4: The Geosphere

As illustrated in Figure 8.5, the **geosphere** is the solid Earth (which sometimes is not so solid when earthquakes or volcanic eruptions occur). The geosphere is an enormous source of natural capital. It provides the platform upon which most food is grown and is the source of plant fertilizers, construction materials, and fossil fuels that humans use. As part of its natural capital, the geosphere receives large quantities of consumer and industrial wastes. Past and current practices of using the geosphere as the anthrosphere's waste dump are ultimately unsustainable. As shown in Figure 8.5, the geosphere interacts strongly with the hydrosphere, atmosphere, biosphere, and anthrosphere. Managing and preserving Earth's natural capital are of utmost importance to sustainability.

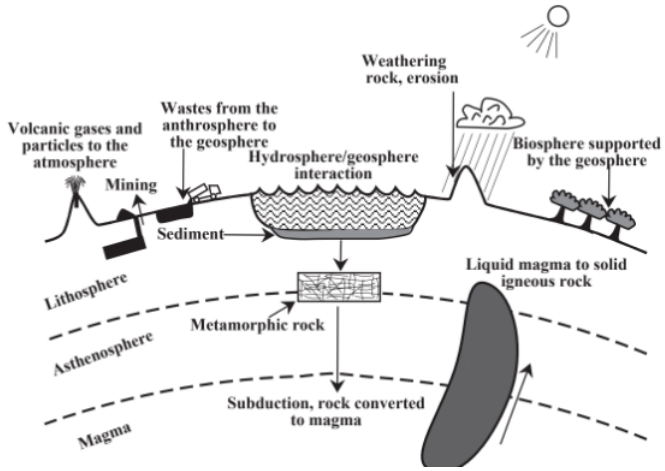


Figure 8.5. The geosphere. The "solid earth" consists of a surface layer of rock about 100 km thick on average floating on top of molten magma. Rock cycles occur within the geosphere. It is an enormous source of natural capital including essential minerals and provides a thin surface of soil upon which most food is grown

Earth is in the shape of a geoid defined by the levels of the oceans and a hypothetical sea level beneath the continents. It is slightly pear-shaped rather than being a perfect sphere because of differences in gravitational attraction in different parts of Earth. Although humans have flown hundreds of thousands of km into space, they have been unable to penetrate more than a few km into Earth's crust.

The geosphere is a layered structure most of which is hot enough to melt rock. Earth's core is a huge ball of iron at a temperature above the normal melting point of iron, but solid due to the enormous pressure that it is under. Above this core is the **mantle** composed of rock and ranging in depth between 300 km and 1,890 km. The deeper inner mantle, though hot enough for the rock to be liquid under ordinary pressures is solid because of the enormous pressure to which it is subjected. On top of the inner mantle is the outer mantle at a depth between 10 km and 300 km composed of hot molten rock called **magma**. Floating on the magma is the solid **lithosphere** composed of relatively strong rock, varying in thickness from just a few to as much as 400 km, averaging about 100 km. The transition layer between the molten magma and the lithosphere is the **athenosphere** composed of hot rock that is relatively weak and plastic. Earth's **crust** is the outer layer of the lithosphere, which is only 5-40 km thick.

Introduced to much controversy in the mid-1900s, the theory of **platetectonics** views Earth'ssurface as consisting of huge lithospheric plates upon which the continents and Pacific Ocean rest, behaving as units. Earth's crust is a dynamic system in which the lithospheric plates move relative to each other by, typically, a few centimeters per year. When abrupt plate movement occurs, an earthquake results. Magma coming to the surface along plate boundaries results in emissions of hot and molten rock, ash, and gases in the form of **volcanoes**.

There are three kinds of boundaries between tectonic plates

- **Divergent boundaries** on ocean floors between tectonic plates that are moving apart are where hot magma undergoes upwelling and cooling to form new solid lithospheric rock, creating ocean ridges.
- **Convergent boundaries** where plates move toward each other forcing matter downward into the asthenosphere in subduction zones, eventually to form new molten magma and in some cases forcing matter upward to produce mountain ranges

- **Transform fault boundaries** where two plates move laterally relative to each other creating fault lines along which earthquakes occur

The conditions outlined above drive the **tectonic cycle**. In this cycle, there is upwelling of molten rock magma at the boundaries of divergent plates. This magma cools and forms new solid lithospheric material. At convergent boundaries, solid rock is forced downward and melts from the enormous pressures and contact with hot magma at great depths, reforming magma. This cycle and the science of plate tectonics explain once puzzling observations of geospheric phenomena including the opening and spreading of ocean floors that create and enlarge oceans, the movement of continents relative to each other, formation of mountain ranges, volcanic activity, and earthquakes.

Earth's Composition

Characterized by definite chemical composition and crystal structure, about 2000 known minerals compose Earth. Most rocks in the crust are composed of only about 25 common minerals. With a composition of 49.5% oxygen and 25.7% silicon, Earth's crust is largely composed of chemical compounds of these two elements with smaller amounts of aluminum, iron, carbon, sulfur, and other elements. Other than aluminum and iron, only about 1.6% of Earth'scrust consists of the kinds of rock that must serve as important resources of metals, phosphorus required for plant growth, and other essential minerals. Careful management of this resource of essential minerals is one of the primary requirements for sustainability.

As molten magma penetrates to near the top of Earth's crust then cools and solidifies, it forms igneous rock. Exposed to water and the atmosphere, igneous rock undergoes physical and chemical changes in a process called **weathering**. Weathered rock material carried by water and deposited as sediment layers may be compressed to produce secondary minerals, of which clays are an important example. Molded and heated to high temperatures to make pottery, brick, and other materials, clays were one of the first raw materials used by humans and are still widely used today.

A part of the crust crucial for the existence of humans and most other non-aquatic life forms is the thin layer of weathered rock, partially decayed organic matter, air spaces, and water composing **soil** (see Chapter 11) that supports plant life. Were Earth the size of a geography classroom globe, the average thickness of the soil layer on it would be only about the size of a human cell! The to player of soil that is most productive of plants is topsoil, which is often only a few centimeters thick in many locations, or even non-existent where poor cultivation practices and adverse climatic conditions have led to its loss by wind and water erosion. The conservation of soil and enhancement of soil productivity are key aspects of sustainability (see Chapter 11).

The Geosphere in Relation to Other Environmental Spheres

Virtually all things and creatures commonly regarded as parts of Earth's environment are located on, in, or just above the geosphere. Major segments of the hydrosphere including the oceans, rivers, and lakes rest on the geosphere, and groundwater exists in aquifers underground. Water dissolves minerals from the geosphere that nourish aquatic life. These minerals and rock particles eroded by moving water from the geosphere are deposited in layers and transformed into rock again. The atmosphere exchanges gases with the geosphere. For example organic carbon produced by photosynthetic plants from atmospheric carbon dioxide may end up as soil organic matter in the geosphere, and the photosynthetic processes of plants growing on the geosphere put elemental oxygen back into the atmosphere. The majority of biomass of organisms in the biosphere is located on or just below the surface of the geosphere. Most structures that are parts of the anthrosphere are located on the geosphere, and a variety of wastes from human activities are discarded to the geosphere. Modifications and alterations of the geosphere can have important effects on other environmental spheres and vice versa.

This page titled 2.5.8.4: The Geosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **8.4: The Geosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.8.5: The Biosphere

Composed of living organisms and the materials and structures that they produce, the **biosphere** is one of the five major environmental spheres. The nature of the biosphere and its essential role in sustainability are the topics of Chapter 13. **Biochemistry**, the chemistry that occurs in the biosphere, is discussed in Chapter 7. The biosphere is of obvious importance to the discussion of soil and agriculture in Chapter 12. Aspects of the biosphere are important in sustainability covered in other chapters of the book.

The living organisms that compose the biosphere are the topic of the science of **biology**. The classes of biomolecules that make up these organisms are outlined in Chapter 7. They include proteins that are the basic building blocks of organisms, carbohydrates made by photosynthesis and metabolized by organisms for energy, lipids (fats and oils), and the all-important nucleic acids(DNA, RNA), genetic materials which define the essence of each individual organism and act as codes to direct protein biosynthesis and reproduction. Hierarchical organization is a characteristic of living organisms in the progression of biomolecules < living cells < organs < organisms < the biosphere, itself. Organisms carry out metabolic processes in which they chemically alter substances to obtain energy and to synthesize new biomass. An essential function of organisms is reproduction, and their young undergo various stages of development. Through their DNA organisms express heredity, and modifications of DNA cause mutations.

Although an imperfect system of classification, organisms are regarded as belonging to several kingdoms. Three of these are organisms capable of existing as single cells, but often occurring in colonies of undifferentiated cells clumped together: (1) Archaebacteria without defined cell nuclei that do not require oxygen or light and often exist in extreme environments such as hot springs; (2) eubacteria without defined cell nuclei including heterotrophs that metabolize organic material, cyanobacteria that obtain energy through photosynthesis, and members that obtain their energy by mediating reactions of inorganic matter; and (3) protists consisting of single-celled organisms that have defined cell nuclei enclosed by a nuclear membrane and often have animal-like features, such as moveable hair-like flagella that enable the organisms to move in water. At more complex levels are generally multicelled plantae (plants) and animalia (animals), as well as fungi including yeasts, molds, and mushrooms.

Several terms are used to describe organisms and their place in ecosystems. A **population** of organisms consists of a group of the same species. Groups coexisting in the same location makeup a **community**. Interacting communities and their physical environment make up an **ecosystem**, all of which grouped together constitute the entire biosphere. The basis of any ecosystem is its **productivity**, the ability to produce biomass, usually through photosynthesis in which organisms remove carbon dioxide from the atmosphere and fix it in the form of organic matter that is further converted by biochemical processes to proteins, fats, DNA, and other life molecules. These biomolecules constitute the basis of the whole ecosystem food chain upon which the remainder of the organisms in the food chain depend for their existence. In water, the most biomass at the base of the food chain is produced by algae, photosynthetic phytoplankton growing suspended in water. Some protozoa and some bacteria also have photosynthetic capabilities. The major photosynthetic organisms in terrestrial ecosystems are plants growing in soil.

The outline of the main features of the biosphere are shown in Figure 8.6. For the most part the biosphere is anchored by dominant plant species that are the major producers of biomass. The dominant plant species may also modify the physical environment in ways that facilitate the existence of other species. The big trees dominant in a rain forest obviously form an environment to which other organisms adapt, for example, by providing safe nesting places for birds. The shade provided by the trees makes up a microclimate and a degree of shelter at ground level in which certain kinds of organisms can thrive. Lichen, a synergistic combination of fungi and photosynthetic algae that grow on the surface of rocks weather the rocks to eventually produce soil in which other plants can grow.

The biosphere has undergone massive evolutionary and climate-related changes over millions of years. Much more rapid and sometimes dramatic changes have been caused by human influences. Arguably the most notable of these took place after Columbus discovered the Americas in 1492. Separated by vast oceans, the biospheres of the Eastern and Western hemispheres had evolved largely independent of each other. As humans introduced organisms from one hemisphere to the other an often spectacular phenomenon called ecological release took place as populations of some species exploded when they were introduced into regions free of their natural predators. The Bluegrass State of Kentucky got its name from the prolific growth of this grass introduced from Europe; clover from Europe also grew rapidly in the New World. Newly introduced peachtrees grew so well in the Carolinas and Georgia that fears were expressed over the potential development of a “peach tree wilderness.” Some regions of Peru were inundated with newly introduced mint. It is possible that a human population explosion in parts of Africa was made possible by

productive corn from the Western Hemisphere enabling the removal of millions of people from the continent into slavery without depleting Africa's population. Originating in the Andes Mountains of South America, the potato became a staple of the diet in Ireland. When the catastrophic *Phytophthorainfestans* fungal blight decimated the crop in the 1840s, Ireland lost half its population to starvation and emigration. Tragically, smallpox introduced to North America by Europeans decimated Native American populations, reducing populations of some tribes by 90%.

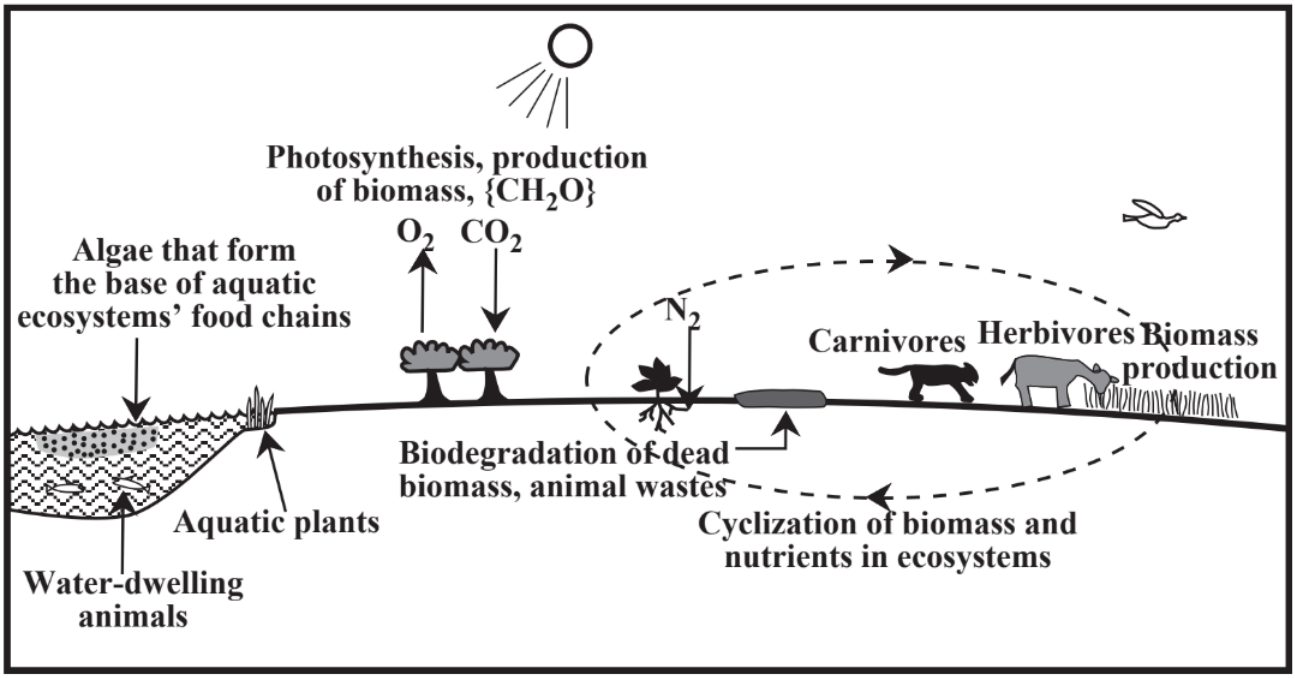


Figure 8.6. Organisms occur within ecosystems in the biosphere where they interact with each other and their enviroment. Ecosystems have evolved in a manner that enables the most efficient utilization of energy and matter with complete recycling of biomass and nutrients. Ecosystems are based upon photosynthetic plants and algae that produce the biomass that anchors the ecosystems, in the process removing carbon dioxide from the atmosphere and returning oxygen. Nitrogen-fixing bacteria, such as those that grow on the roots of leguminous plants, convert elemental nitrogen from the atmosphere to biochemically bound nitrogen. Fungi and bacteria degrade dead biomass, a process that returns nutrients to the ecosystem.

The biosphere has significant effects on the other environmental spheres and *vice versa*. Materials generated in the biosphere are used in the anthrosphere; wood is one such material. Biological productivity is largely determined by conditions of geospheric soil. The availability of the hydrosphere's water largely determines kinds and populations of organisms. Fertilizers and pesticides produced in the anthrosphere have a strong influence on biological productivity in agriculture. Shielding organisms from pollutants, wastes, and toxic substances generated in the anthrosphere is a high priority in environmental protection. The design and operation of the anthrosphere strongly influence the nature of the biosphere and its productivity, especially in the agricultural sector.

The biosphere largely determines Earth's environment. The atmosphere's oxygen was generated by photosynthetic bacteria eons ago. Lichen communities of synergistically growing algae and fungi act upon geospheric rock forming soil that supports plant life. The anthrosphere of less industrialized societies has been largely a product of the biosphere in which the human residents have existed. Massive herds of bison provided food, robes, and the material for the teepee dwellings of North American Plains Native Americans. The availability and kinds of wood have largely determined the nature of dwellings constructed in many societies. More than 2000 years ago, domesticated animals harnessed to carts, wagons, and plows provided humans with mobility and the source of power used to cultivate soil. Many societies, including the Amish farmers in the U.S., still use horses, donkeys, mules, oxen, and water buffalo for land cultivation and transport of goods. One of the basic tenets of green chemistry and sustainability is the use of materials from the biosphere to replace those produced from scarce, expensive petroleum.

In achieving sustainability humans have much to learn from the organisms in the biosphere which over millions of years have developed essential tools of survivability. An important aspect is the ability to thrive and be productive under the mild, safe conditions under which organisms may exist, conditions which are most desirable for green chemical synthesis and other sustainable activities. Even the conditions under which some thermophilic bacteria thrive in the thermal hot springs of Yellowstone

National Park are mild compared to the much higher temperature, high-pressure conditions required in many chemical syntheses. The intolerance of living organisms to toxic substances provides valuable lessons regarding which substances should be avoided in the practice of green chemistry.

Important lessons in the development of sustainable systems of industrial ecology (Chapter13) are provided by the biosphere. Over hundreds of millions of years of evolution organisms in the biosphere have had to evolve sustainable ecosystems for their own survival, completely recycling materials and enhancing their environment. In contrast, humans have behaved in ways that are unsustainable with respect to their own existence, exploiting nonrenewable resources and fouling the environment upon which they depend for their own survival. The complex, sustainable ecosystems in which organisms live sustainably in relationships with each other and their surroundings serve as models for anthropogenic systems. By taking lessons from the biosphere and its long-established ecosystems, humans can develop much more sustainable systems of industry and commerce.

A crucial respect in which the biosphere is a key to achieving sustainability is its ability to perform photosynthesis and synthesis of specialized materials. Using carbon dioxide from the atmosphere and energy from the sun, organisms produce biogenic materials in a much greener, safer, more sustainable manner than the manner in which materials are produced in the anthrosphere. Furthermore, organisms are particularly well adapted to make a variety of complex and specialized materials that are very difficult or impossible to make by purely chemical means.

This page titled 2.5.8.5: The Biosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **8.5: The Biosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.8.6: The Anthrosphere

Although some authorities object to recognizing the things that people make, modify, and use as a separate environmental sphere — the **anthrosphere**—as a distinct environmental sphere, it is essential to consider it in the achievement of sustainability. Just a look around us shows the dwellings, buildings, roads, airports, factories, power lines and numerous other things constructed and operated by humans as visible evidence of the existence of the anthrosphere on Earth (see Figure 8.7). The anthrosphere and its influences are so obvious and even intrusive that the Nobel

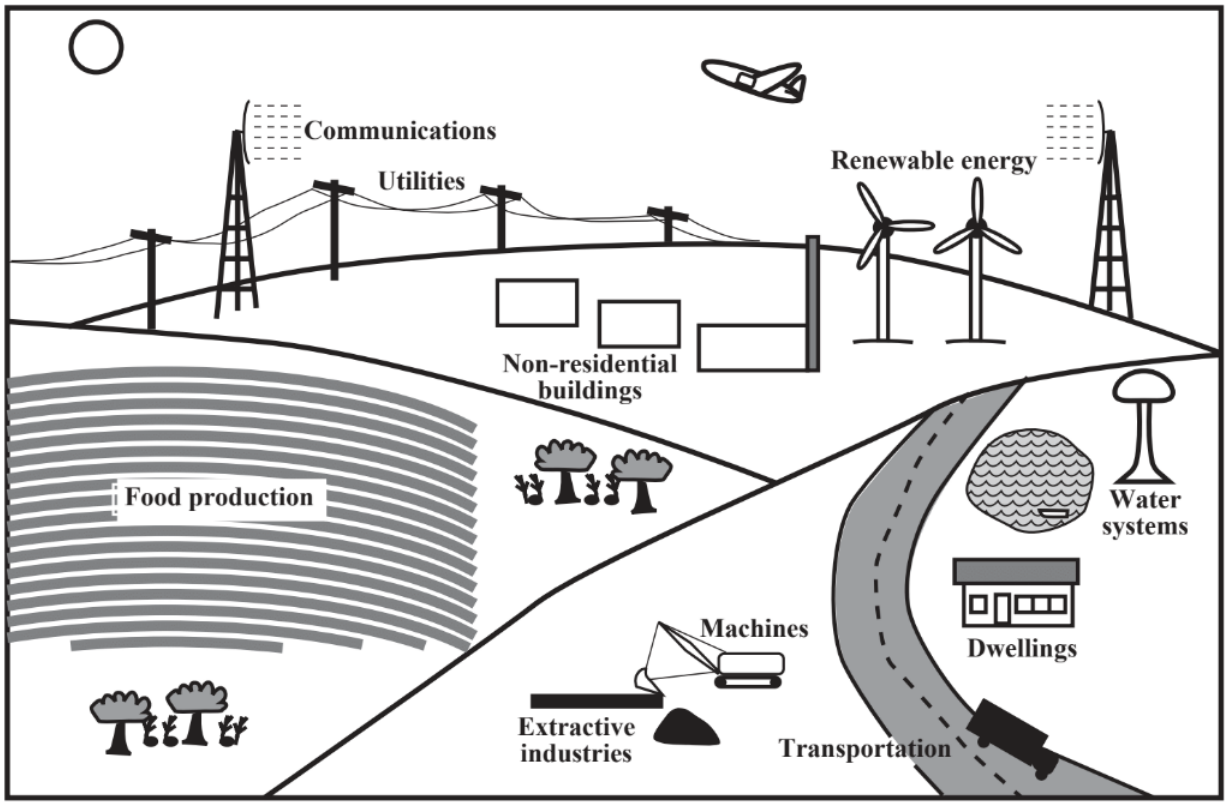


Figure 8.7. There are many aspects to the anthrosphere as illustrated by the examples in this figure. It is tied closely with the other environmental spheres

Prize-winning atmospheric chemist Paul Crutzen has argued convincingly that Earth is undergoing a transition from the holocene geological epoch to a new one, the **anthropocene**. This is occurring because human activities are now quite significant compared to nature in their impact on Earth'senvironment and are changing Earth's fundamental physics, chemistry, and biology. There is concern that, especially through changes in global climate, activity in the anthrosphere will detrimentally alter Earth's relatively stable, nurturing environment and produce one that is much more challenging to human existence.

It is not completely accurate to believe that pre-industrial humans had little influence on the biosphere and ecosystems. The grazing habits of excessive numbers of goats bred by humans form eat and milk centuries ago contributed to the transformation of formerly productive grasslands of the Middle East and North Africa to deserts. Some pre-Columbian native South Americans created their unique anthrosphere by building terraces, draining wetlands, and constructing raised agricultural fields to grow a variety of food crops. As much as one third of the Midwestern prairies first encountered by early European settlers in the present day U.S. were created through the pyromaniacal tendencies of Native Americans who burned forests to create grasslands that supported game animals. Subsequently, forestation, largely with European tree varieties, often followed in the paths of the first non-native settlers. Compared to the modern industrial era, however, the impact of humans until just a few centuries ago was comparatively benign. Particularly with the development of fossil fuel resources and the large machines powered by these fuels, during the last approximately 200 years, humans have built a pervasive anthrosphere that is having massive environmental impact.

The boundaries between the anthrosphere and the other environmental spheres are sometimes blurred. Most of the anthrosphere including buildings, highways, and railroads is anchored to the geosphere. Gardens that adorn the anthrosphere are planted on geospheric soil and the flowering plants in them are part of the biosphere. Farm fields are modifications of the geosphere, but the crops raised on them are part of the biosphere. Coal mines are burrowed into the anthrosphere. Ships move over ocean waters in the hydrosphere and airplanes fly through the atmosphere.

There are many distinct segments of the anthrosphere as determined by a number of factors including where and how humans dwell; the movement and distribution of goods; the provision of services; the utilization of non-renewable materials; the provision of renewable food, fiber, and wood; the collection, conversion, and distribution of energy; and the collection, treatment, and disposal of wastes. With these factors in mind it is possible to list a number of specific things that are parts of the anthrosphere as shown in Figure 8.7. These include dwellings as well as other structures used for manufacturing, commerce, education, and government functions. Utilities include facilities for the distribution of water, electricity, and fuel, systems for the collection and disposal of municipal wastes and wastewater (sewers), and — of particular importance to sustainability, — systems for materials recycle. Transportation systems include roads, railroads, and airports, as well as waterways constructed or modified for transport on water. The anthrospheric segments used in food production include cultivated fields for growing crops and water systems for irrigation. A variety of machines, including automobiles, trains, construction machinery, and airplanes are part of the anthrosphere. The communications sector of the anthrosphere includes radio transmitter towers, satellite dishes, and fiber optics networks. Oil and gas wells are employed for extracting fuels from the geosphere and mines are excavated into the geosphere for removing coal and minerals.

The Crucial Infrastructure

A critical part of the anthrosphere may be classified as **infrastructure**. Infrastructure is generally considered to be parts of the anthrosphere used by large numbers of people in common. It consists of utilities, facilities and systems essential to a properly operating society. Physical components of the infrastructure include electrical power generating facilities and distribution grids, communications systems, roads, railroads, air transport systems, airports, buildings, water supply and distribution systems, and waste collection and disposal systems. A very important part of infrastructure is non-physical, composed of laws, regulations, instructions, and operational procedures. Because they are used in common by many people, crucial parts of the infrastructure are in the public sector; other segments are privately owned and operated. For example, major airports are almost always publicly owned whereas the aircraft that serve them are generally owned by private corporations.

Consider a computer operating system as infrastructure. The use of the system enables a computer to run programs for word processing, record keeping, calculation, drawing, communication, and other common computer operations. The operating system enables the computer to properly record, store, correlate, and output the products of the programs that it operates. By analogy, the infrastructure of the anthrosphere facilitates infrastructure activities including acquisition and processing of materials, the conversion of materials to manufactured items, and the distribution of such items. Crash-prone computer systems that are outdated or poorly designed in the first place cause lost productivity and general distress to the computer user. Similarly, an outdated, cumbersome, poorly designed, worn-out anthrospheric infrastructure causes economic systems and societies to operate in a very inefficient manner that is inconsistent with sustainability. Catastrophic failure can result as has occurred with cascading breakdowns of electrical power grids or failure of wastewater treatment systems resulting in discharge of sewage to streams.

Deterioration of the infrastructure is a continuing concern. One of the greatest problems is corrosion, a chemical process in which metals, such as the steel that composes bridge girders, tend to revert to the state in which they occur in nature (in the case of steel, rust). Human negligence, misuse and vandalism can all cause premature loss of infrastructure function. A major concern with terrorism is potential damage to infrastructure including cybercrime that could cripple electrical distribution systems. Infrastructure problems frequently begin with improper design. Sustainability requires that elements of the infrastructure be properly designed, maintained, and protected to avoid the expense and material and energy required to rebuild infrastructure if it fails before its expected lifetime is up.

To date much of infrastructure has been dedicating to thwarting what nature does naturally, often a losing proposition in the long run. A prime example consists of levees constructed along rivers which work well until overwhelmed by extraordinarily heavy precipitation events. Much more consideration must be given to sustainability and the maintenance of environmental quality in the development of infrastructure. Examples include highly effective waste treatment systems with recovery of materials and energy

from wastes, high-speed rail systems to replace inefficient movement of people and freight by private carriers, and electrical systems that use wind power to the maximum extent possible.

The Sociosphere

The **sociosphere** is the societal organization of people including their governments, laws, cultures, religions, families, and social traditions and is a critical part of infrastructure. A well functioning sociosphere enables people to lead good lives within a sustainable environment and economic system. Largely because of disfunctional social systems, the quality of life and the environment in some countries with substantial resources, especially of petroleum, is often sub-standard. Societies in countries with dictatorial, corrupt governments that do not nurture human rights are not beneficial in the maintenance of sustainability. Sustainability and quality of life are also not well served by anti-government creeds that reject the role of well-functioning governments in implementing sensible well-administered laws and regulations designed to protect the environment and maintain sustainability.

An important consideration in the sociosphere is the science of **economics**, which describes the production, distribution, and use of income, wealth, and materials (commodities). Much of economics as it is usually practiced is inconsistent with the development of sustainability upon which functional economic systems must ultimately depend. Economic value has traditionally been measured in terms of financial and material possessions with emphasis on growth and with an arrow view of the environment. Earth has been largely regarded as a part of the economic system from which materials may be extracted, which is to be "developed" with structures and other artifacts of the anthrosphere, and into which wastes are to be discarded. Such an approach is putting an unacceptable strain on environmental support systems and Earth's natural capital. Amore enlightened economic view regards Earth's natural capital as an endowment. As with financial endowments, Earth's endowment of natural capital should be nurtured, with only a portion of its income spent for immediate needs and the rest devoted to enhancing the natural capital. Therefore, it is essential for sustainability that economics be viewed as a part of Earth'sgreater environmental system, rather than viewing the environment as a subsection of a world economic system. Instead of defining wealth in material possessions it should be measured in terms of well-being and satisfaction with life, operating within rules that promote and require sustainability.

The Human Microsociosphere

The immediate environment in which humans carry out their daily activities may be regarded as a human microsociosphere. For many people recent years have seen a marked change in this microenvironment as the consequence of a flood of electronic devices including personal computers, cellular telephones, and other devices that flood individuals with information and communications. Some psychological studies have suggested that this deluge of information is changing human behavior, not necessarily in beneficial ways. Many people engage in multitasking, for example by texting or talking on cell phones while working or performing other tasks. The consequences of this can even be deadly as happens when people get into automobile accidents while talking or — worst of all — texting on a cell phone. Some people have several computers and screens at their workplace and attempt to follow several streams of information simultaneously. It is the feeling of some reputable neuroscientists that this kind of activity forces the human brain to function in ways that it has not evolved to handle making multitaskers more stressed, impatient, impulsive, and forgetful. Extreme cases have been documented of people who fall asleep with a laptop computer on their chest and whose first action after waking up is to connect to the internet.

Despite the benefits of modern computer and communications technology, the "always-plugged-in existence" that many people now lead has a definite downside. Not the least of these harmful effects can be upon human social interaction in which the major kind of interaction is through a computer rather than with other human beings. On the other hand, e-mail and other modes of electronic communication have made it possible for people to stay in constant, virtually instantaneous contact and to expand their circle of human contact, even if it is over a computer screen. The challenge of dealing with the kinds of problems discussed here is expected to be most acute for the increasing numbers of people who work from home on computers.

This page titled 2.5.8.6: The Anthrosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **8.6: The Anthrosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.8.7: Cycles of Matter

The physical connections among the environmental spheres are largely through cycles of matter. Involving physical processes, chemical reactions, and biochemical processes, they are called **biogeochemical cycles**. They are commonly named for the major element involved in each, usually an essential organism nutrient including carbon, nitrogen, oxygen, sulfur, and phosphorus. A particularly important cycle is the hydrologic cycle of water shown in Figure 8.1. In Figure 8.5 aspects of the **rockcycle** are shown in which molten rock solidifies, undergoes weathering, maybe carried by water and deposited as sedimentary rock, is converted to metamorphic rock by heat and pressure, and is eventually buried at great depths and melted to produce molten rock again. Anthrospheric processes are also very much involved in the important cycles of matter such as the input of chemically fixed nitrogen synthesized from atmospheric N_2 in the nitrogen cycle.

The two major parts of biogeochemical cycles are the **reservoirs** in which matter is contained for some time and the **conduits** through which matter and energy are moved. Oceans, the atmosphere, parts of Earth's crust, and organisms are reservoirs of matter. The atmosphere acts as a conduit to carry water vapor in the hydrologic cycle and streams carry sedimentary matter in the rock cycle. **Endogenic cycles**, of which the phosphorus cycle is an example, are those that occur below or directly on the surface of the geosphere without a significant atmospheric component; most biogeochemical cycles are **exogenic** cycles in which the atmosphere serves as a conduit and often as a reservoir.

Figure 8.8 illustrates one of the key biogeochemical cycles, the carbon cycle. An important reservoir of carbon is in the atmosphere as carbon dioxide. Photosynthetic processes by plants extract significant amounts of carbon from the atmosphere and fix it as biological carbon in the biosphere. In turn, animals and other organisms in the biosphere release carbon dioxide back to the atmosphere through the respiration processes by which they utilize oxygen and food for energy production. More carbon dioxide is released to the atmosphere by the combustion of biological materials such as wood and fossil fuels including coal and petroleum. Carbon dioxide from the atmosphere dissolves in water to produce dissolved carbon dioxide and inorganic carbonates. Solid carbonates, particularly limestone, dissolve in bodies of water to also produce dissolved inorganic carbon. Respiration of organic matter by organisms in water and sediments also

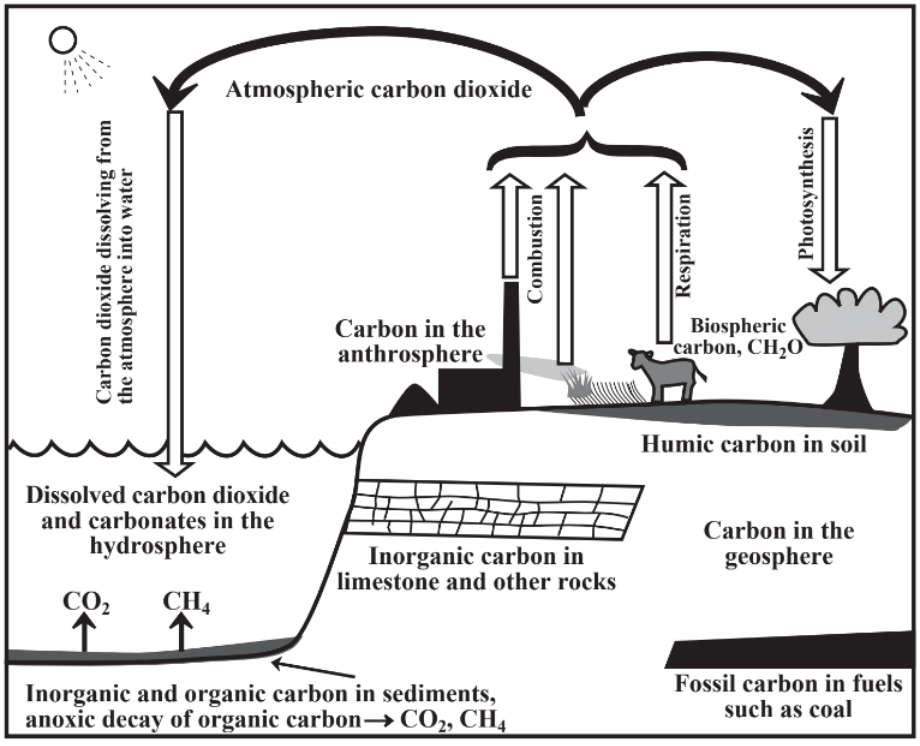


Figure 8.8. The carbon cycle showing the various reservoirs and conduits of carbon species in the environment. Because some important carbon compounds contain oxygen and elemental oxygen is released to the atmosphere by the photosynthetic process that fixes carbon from atmospheric carbon as biomass, the carbon cycle is closely related to the oxygen cycle.

produces dissolved inorganic carbon species in water. Large amounts of carbon are held in the geosphere. The major forms of this carbon are fossil carbon of ancient plant origin held by fossil fuels, such as coal and petroleum, and inorganic carbon in carbonate rocks, especially limestone. Organic carbon from partially degraded plant material is present as humic material in soil, a source that is added to by organisms breaking down plant biomass from the biosphere.

The carbon cycle is extraordinarily important in maintaining sustainability because a major part of it is the fixation of carbon from highly dilute atmospheric carbon dioxide into biomass by photosynthesis carried out by green plants. Biomass is a source of food, chemical energy, and raw materials and the carbon cycle contains the main pathway by which solar energy is captured and converted to a form of energy that can be utilized by organisms and as fuel.

Other important cycles of matter are linked to the carbon cycle. The oxygen cycle describes movement of oxygen in various chemical forms through the five environmental spheres. At 21% elemental oxygen by volume, the atmosphere is a vast reservoir of this element. This oxygen becomes chemically bound as carbon dioxide by respiration processes of organisms and by combustion. The reservoir of atmospheric oxygen is added to by photosynthesis. Oxygen is a component of biomass in the biosphere and most rocks in Earth's crust are composed of oxygen-containing compounds. With its chemical formula of H_2O, water in the hydrosphere is predominantly oxygen.

In addition to the carbon and oxygen cycles described above, three other important life-element cycles are those of nitrogen, sulfur, and phosphorus:

•**Nitrogen cycle**: Biochemically bound nitrogen is essential for life molecules including proteins and nucleic acids. Although the atmosphere is about 80% by volume elemental N_2, this molecule is so stable that it is difficult to split it apart so that N can combine with other elements. This process is performed in the anthrosphere by the synthesis of NH_3, from N_2 and H_2 over a catalyst at high temperatures and very high pressures. Furthermore, air pollutant NO and NO_2 form from the reaction of N_2 and O_2 under the extreme conditions in internal combustion engines. In contrast, some bacteria, including *Rhizobium* bacteria growing on the roots of legume plants, convert atmospheric nitrogen to nitrogen compounds under the very mild conditions just below the soil surface. Plants convert nitrogen in NH_4^+ and NO_3^- to biochemically bound N. As part of the nitrogen cycle, biochemically bound nitrogen is released as NH_4^+ by the biodegradation of organic compounds. The nitrogen cycle is completed by microorganisms that use NO_3^- as a substitute for O_2 in energy-yielding metabolic processes and release molecular N_2 gas to the atmosphere. Other than nitrogen fixation in the anthrosphere and formation of nitrogen oxides in the atmosphere from lightning discharges, most transitions in the nitrogen cycle are carried out by organisms, especially microorganisms.

•**Sulfur cycle**: The sulfur cycle includes both chemical and biochemical processes and involves all spheres of the environment. Chemically combined sulfur enters the atmosphere as pollutant H_2S and SO_2 gases, which are also emitted by natural sources including volcanoes. Large quantities of H_2S are produced by anoxic microorganisms degrading organic sulfur compounds and using sulfate, SO_4^{2-}, as an oxidizing agent and discharged to the atmosphere. Globally, a major flux of sulfur to the atmosphere is in the form of volatile dimethyl sulfide, (CH_3S), produced by marine microorganisms. The major atmospheric pollutant sulfur compound is SO_2 released in the combustion of sulfur-containing fuels, especially coal. In the atmosphere, gaseous sulfur compounds are oxidized to sulfate, largely in the forms of H_2SO_4 (pollutant acid rain) and corrosive ammonium salts (NH_4HSO_4) which settle from the atmosphere or are washed out with precipitation. The geosphere is a vast reservoir of sulfur minerals including sulfate salts ($CaSO_4$), sulfide salts (FeS), and even elemental sulfur. Sulfur is a relatively minor constituent of biomolecules, occurring in two essential amino acids, but various sulfur compounds are processed by oxidation-reduction biochemical reactions of microorganisms.

•**Phosphorus cycle**: Unlike all the exogenous cycles with an atmospheric component discussed above, the phosphorus cycle is endogenous with no significant participation in the atmosphere. It is an essential life element and ingredient of DNA as well as ATP and ADP through which energy is transferred in organisms. Dissolved phosphate in the hydrosphere is an essential nutrient for aquatic organisms, although excessive phosphate may result in too much algal growth causing an unhealthy condition called eutrophication. Phosphorus is abundant in the geosphere, especially as the mineral hydroxyapatite, chemical formula $Ca_5OH(PO_4)_3$. Significant deposits of phosphorus-rich material have be enformed from the feces of birds and bats (guano).

This page titled 2.5.8.7: Cycles of Matter is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **8.7: Cycles of Matter** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

https://bio.libretexts.org/@go/page/108839

2.5.8.8: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1.Observations such as heavy clouds, sometimes torrential rains, and humid air might lead to the conclusion that the atmosphere is a major reservoir of water. Using internet resources, find out if this is true and compare the amount of water in Earth's atmosphere to that in the oceans.

2.Suggest a role played by the tropopause in the hydrologic cycle. Without the tropopause, would Earth even have a hydrologic cycle?

3. Water of the Colorado River of the U.S. is essentially fully committed for irrigation and municipal water supplies. What fraction of the Mississippi River discharge would be equivalent to the amount of water used from the Colorado River?

4. Argon and neon are byproducts of the production of another important industrial product. What is the main process of which these two gases are byproducts?

5. It is stated in this chapter that if the stratospheric ozone layer were in a pure layer of ozone at Earth's surface, the layer would be only about 3 millimeters thick. According to the ideal gas law, a mole of gas occupies 22.4 liters at a temperature of 273 K and a pressure of 1 atmosphere, very close to average conditions at sea level. Using this information calculate a reasonable estimate of the total mass of ozone in the ozone layer.

6. Although the ozone layer mentioned in the preceding question is a very good thing and even essential for human life, where in the atmosphere is ozone detrimental? Explain.

7. Long before humans were able to reach the atmospheric region of the ionosphere scientists observed the reception of radio waves transmitting over long distances far beyond line-of-sight, a "skipping" effect that enabled (at times, with luck) short wave transmissions across the Atlantic. The hypothesis was that the signals were bouncing off a layer of ions many kilometers up in the atmosphere. They also observed that longer distance transmission was possible at night. From this information suggest how ions are formed in the ionosphere and why it raises during nightime.

8. What is a common air pollutant that is associated with the sulfur cycle?

9. Some large volcanic eruptions cause significant global cooling whereas the climatic effects of other, equally large eruptions are negligible. Very little, if any of the cooling is caused by volcanic ash ejected to the atmosphere, large as the amounts of ash may be. Give a plausible explanation for these observations.

10. Soil is formed over very long periods of time beginning with igneous rock. Explain why soil forms faster in wet, temperate climates subjected to seasonal changes than it does in hot, dry regions where it may be largely absent.

11. What lessons are offered by the biosphere for the development of workable systems of industrial ecology?

12. What process in the anthrosphere is most closely analogous to photosynthesis by plants and algae in the biosphere? Explain

13. Consider your computer as a system of industrial ecology. What is its infrastructure? To which parts of an industrial ecosystem is it analogous? Consider how word processing, data, and graphics were handled in the days when typewriter, calculator, and paper were the instruments of choice and suggest how a computer recycles data as compared to those times.

14. Look up the definition of economics, especially if one can be found in older textbooks. Suggest an alternate definition that is consistent with sustainability and the preservation of Earth's natural capital.

15. Since about 1980 the classification of organisms such as the generally recognized kingdoms of organisms has undergone significant change. Suggest a development that has enabled this change to occur.

16. What are plankton? What kind of plankton are especially important for aquatic ecosystems?

17. By around 1800 one form of renewable energy was dominant in the generation of power for applications such as grinding grain for flour and sawing wood, but was not usable during the coldest winter months. Suggest the source of power and why it was not usable during winter.

18. Suggest the major aspects of the silicon cycle. Is it exogenous or endogenous? Does it have a biospheric component?

2.5.8.8: Questions and Problems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** is licensed CC BY-NC-SA 4.0.

2.5.8.9: Supplementary References

Florinsky, Igor V., Ed., *Man and the Geosphere*, Nova Science Publishers, Hauppauge, NY, 2009.

Hyde, Natalie, *What is the Rock Cycle?*, Crabtree Publications, New York, 2011.

Silivanch, Annalise, *Rebuilding America's Infrastructure*, Rosen Publications, New York, 2011.

2.5.8.9: Supplementary References is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Supplementary References** is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

2.5.9: Water - the Ultimate Green Substance

“Although the wastewater from Mexico City that the farmers in Mexico’s Mezquital valley use for irrigating their crops is a foul concoction that sometimes causes boils and other ailments to those who contact it, the farmers that use it are concerned that a new drainage and treatment system may remove nutrients and enable recycle of water back to the city depriving them of the irrigation water upon which their livelihoods depend.

2.5.9.1: H_2O - Simple Formula, Remarkable Molecule

2.5.9.2: Occurrence, Availability , and Utilization of Water on Earth

2.5.9.3: Water Chemistry

2.5.9.4: Water Pollution

2.5.9.5: Greening of Water - Purification Before Use

2.5.9.6: Wastewater Treatment

2.5.9.7: Advanced Wastewater Treatment and Recycling

2.5.9.8: The Many Uses of Water

2.5.9.9: Hot Water - Pressurized Subcritical Water

2.5.9.10: Supercritical Water

2.5.9.11: Literature Cited

2.5.9.12: Questions and Problems

2.5.9.13: Supplementary References

This page titled 2.5.9: Water - the Ultimate Green Substance is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.9.1: H_2O - Simple Formula, Remarkable Molecule

Water composes the hydrosphere, which is described and discussed in Section 8.2. This chapter enlarges upon that discussion and the crucial role of water in the environment. In this chapter and elsewhere in the book the term **naturalwater** is used in reference to water that occurs in the environment in comparison to water in the anthrosphere, such as water in municipal water distribution systems. Water has a special place in living organisms and the environment. The quality and availability of water are of the utmost importance to humans and the environment. Although scarce and badly polluted in many parts of the world, water is arguably the most recyclable of the substances that compose Earth's green capital and it is accurately described as the ultimate green substance.

This chapter addresses major aspects of the sustainability of water. The first of these is water pollution, which degrades water quality and can make it unfit for use or to support life. The second major area addressed is how water quality can be maintained and enhanced, largely through various water treatment processes. A third major area is water pollution prevention and a fourth is recycling of the water resource which is arguably nature's most recyclable material.

The chemical formula of water, H_2O, is probably the best known of all compounds. This simple formula represents a substance that is unique and complex in its behavior. These special properties are due to the molecular structure of the H_2O molecule represented in Figure 9.1. There are four pairs of electrons in the outer electron shell of the O atom in the H_2O molecule, two of which compose the bonds between the H and O atoms and two of which are lone pairs. The distribution of these pairs as far apart as possible around an imaginary sphere representing the outer electron shell of the O atoms results in the two H-O bonds being located at an angle rather than a straight line. The side of the molecule with the two H atoms has a partial positive charge and the side with the two non-bonding pairs has a partial negative charge, so the molecule is polar. This polarity and the ability of the H atoms on one molecule to form hydrogen bonds to O atoms on other molecules determine the remarkable chemical and physical diversity of water.

Especially because of their hydrogen bonding capability, water molecules are strongly attracted to each other. This means that a large amount of heat energy must be put into a mass of water to enable the molecules to move more rapidly as the temperature is raised. This gives water a very high **heat capacity**. A very large amount of energy must be put into a mass of ice to break the hydrogen bonds holding the molecules in place in the solid as it melts and an equally large amount of heat energy is released when liquid water freezes. Thus water has a very high **latent heat of fusion**. Even more energy per unit mass is required to convert liquid water to vapor (steam) and an equal amount of energy is released when water vapor condenses to liquid. This means that water has a very high **heat of vaporization**.

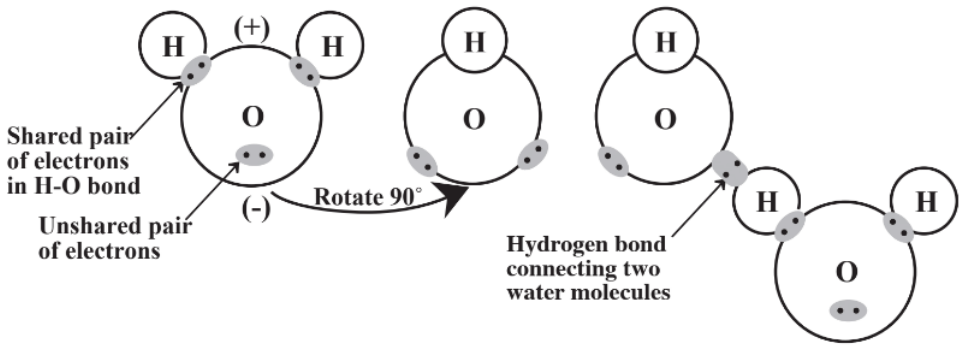


Figure 9.1. Representation of the polar nature and hydrogen bonding capability of the water molecule that give it unique properties that determine its essential role in living organisms and the environment.

The ability of water to absorb, release, and store heat is crucial to its role in the environment and its practical uses.[1] Water's high heat capacity stabilizes temperatures of organisms and geographical areas. Steam produced in a boiler can be transferred through insulated pipes to remote locations and condensed to release heat. The heat released when atmospheric water condenses warms the surrounding air and is the driving force behind tropical storms. Europe owes its relatively mild weather despite its northern latitudes to heat carried by water across the North Atlantic Ocean from the Gulf of Mexico. As the water releases heat and cools along the European coasts, its density increases and it flows at lower ocean depths back to the Gulf of Mexico to repeat the cycle. Water's high latent heat of fusion stabilizes temperatures of bodies of water at water's freezing point (0°C).

In addition to those listed above, there are other unique and environmentally important qualities of water. It is an excellent solvent, especially for ionic substances, making it important in the transport of nutrients and wastes in the biosphere and in the dissolution, transport, and deposition of minerals in the geosphere. Water has a very high surface tension, a controlling factor in physiology and a property involved in formation of drops in rainfall. The temperature/density relationship of water causes bodies of water to become stratified (see Figure 9.4), a property that strongly affects the chemical behavior of water in stratified bodies of water. The fact that the maximum density of water occurs as a liquid at 4°C means that solid ice floats. If that were not the case, bodies of water in northern climates would become frozen solid with only a surface layer thawing during the summer. Water is largely transparent to visible light, which can penetrate the liquid to some depth and enable photosynthetic phytoplankton and some bottom-rooted plants in shallow water regions to carry out photosynthesis and produce the biomass that is the basis of aquatic food chains (see Section 8.5).

This page titled 2.5.9.1: H_2O - Simple Formula, Remarkable Molecule is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **9.1: H_2O - Simple Formula, Remarkable Molecule** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.9.2: Occurrence, Availability , and Utilization of Water on Earth

Much of Earth's approximately 1.4 billion cubic kilometers (1.4×109km3) water that is not chemically bound circulates through the hydrologic cycle represented in the illustration of the hydrosphere shown in Figure 8.1. About 97.6% of this water is salt water in oceans. Of the approximately 33 million km^3 of water remaining, 86.9% is in the solid form in snowpack, ice and glaciers; 12.0% is accessible water under ground (groundwater), 0.37% is in freshwater lakes, reservoirs, and ponds; 0.31% is in saline lakes; 0.19% is soil moisture; 0.19% is in biospheric organisms; 0.039% is in the atmosphere as vapor and cloud droplets; 0.011% is in wetlands; and only 0.0051% constitutes the water in all Earth's rivers, streams, and canals. Therefore, a remarkably small percentage of Earth's water is relatively available for use, a major factor in the sustainability of this crucial part of Earth's natural capital.

Fresh surface water and groundwater are the major sources for human use. Desalinated seawater is a growing source of fresh water in some areas that do not have access to this valuable resource. Most seawater desalination plants are in the Middle East, an arid region with generally abundant sources of energy that can be used for desalination. The world's largest desalination plant is the Jebel Ali multi-stage flash distillation unit in the United Arab Emirates that can produce 300 million cubic meters of potable water per year. There are significant resources of brackish (saline) groundwater throughout the world. In many cases this water has a lower salt content than seawater making it easier and more economical to desalinate.

A major concern regarding water's availability is its uneven distribution with respect to time and location. Some areas experience tremendous rainfall and flooding during wet (monsoon)seasons with dry conditions in between. Whereas these fluctuations are predictable and compensation may be made for them, greater problems occur with long-term droughts. Regions of Africa are periodically afflicted with droughts that last for several years, killing crops and animals and inflicting great hardship on the people in the region. Evidence from tree rings indicates a pre-Columbian drought of almost three centuries duration in what is now the southwestern U.S.! Uneven geographic distribution of water occurs throughout the world and is illustrated for the continental U.S. in Figure 9.2. It is seen that precipitation is generally adequate in the eastern part of the country, although damaging droughts do occur in this region. However, the western continental U.S. has a shortage of precipitation with essentially permanent drought conditions in some regions including Nevada, Arizona, and southern California. The problem is exacerbated by the popularity of these regions as areas in which people want to live and the demands that they put on limited water resources.

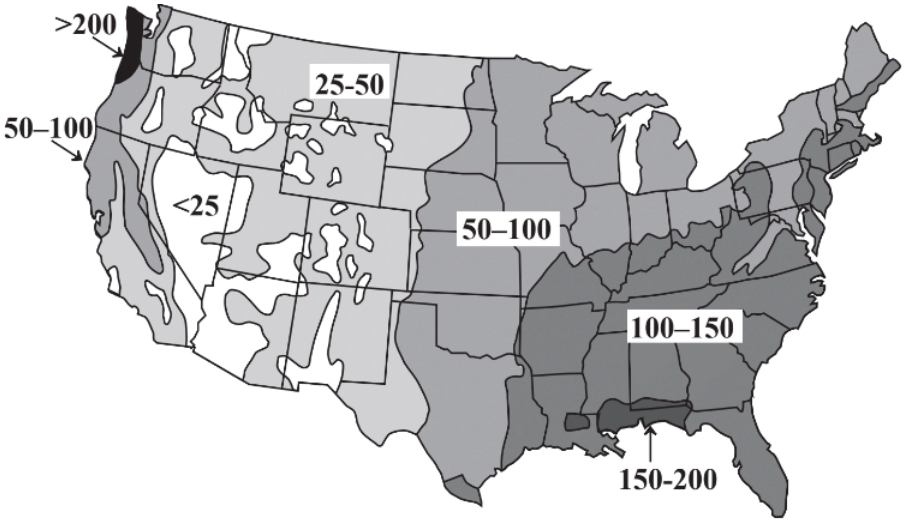


Figure 9.2. Distribution of average annual rainfall in the continental United States, amounts in centimeters per year.

Surface Streams of Water

A major source of water is surface water that flows in streams and rivers. The areas of land upon which precipitation falls to provide flows of surface water is called the **watershed**. Insofar as water utilization is concerned, watersheds constitute one of the most important connections between the geosphere and the hydrosphere; the nature of the watershed largely determines the quantity and quality of water available for human use. A good watershed retains water for a significant length of time, which

reduces flooding, allows for a steady flow of runoff water, and maximizes recharge of water into groundwater reservoirs (aquifer recharge, Figure 9.3). Measures employed to enhance watershed quality include minimization of cultivation and forest cutting, especially on steeply sloping sections of watershed, construction of terraces and waterways planted to grass on cultivated land to minimize erosion and accumulation of sediment, construction of small impoundments on feeder streams of the watershed, and of course minimization of pollution such as from herbicides.

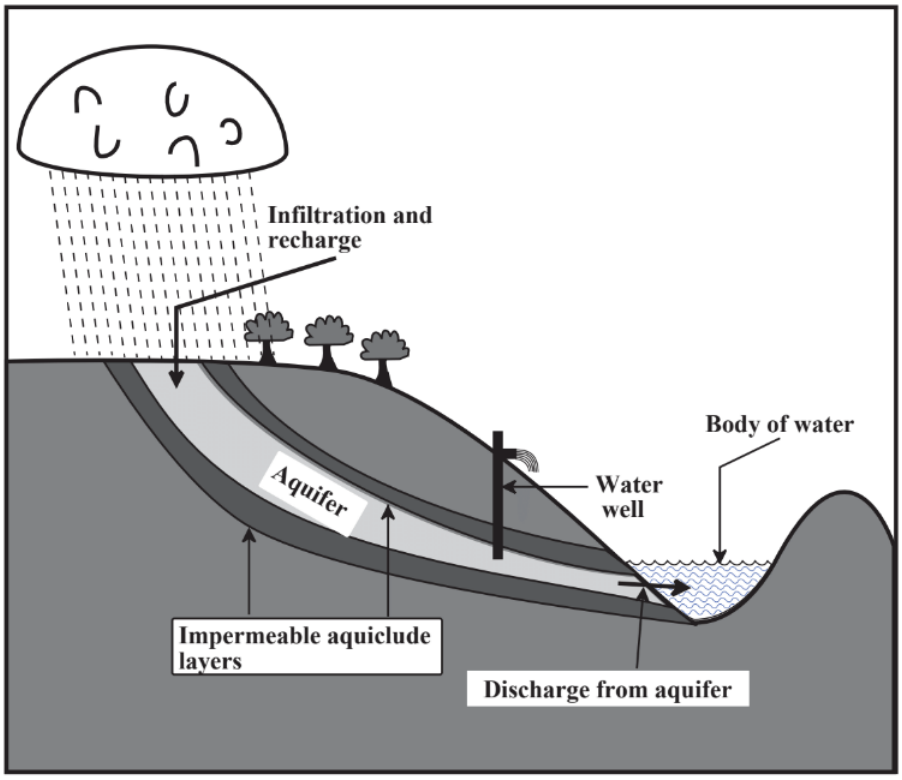


Figure 9.3. Illustration of groundwater in aquifers in the geosphere. Groundwater enters aquifers from surface water in recharge zones and is discharged through springs or pumped out through wells.

An important aspect of streams that affects water utilization and quality is their ability to mobilize sedimentary materials through erosion, transport materials along with stream flow, and deposit them as solids. Normally it is desirable for streams used as a source of water to have minimal sedimentary material, which has to be removed during water treatment.

Most rivers, once free flowing and unimpeded by human intervention, have been modified by humans to generate power, for water supply, and reduce effects of flooding. Beginning with impoundment of water from the Owens Valley in northern California and extending to the Feather, Sacramento, and Colorado rivers, Los Angeles' voracious thirst for water is served by a vast system of dams, canals, and tunnels. Diversion of water from the Owens valley has turned a formerly productive agricultural area into one of limited agricultural use. Many of the adverse effects of water diversion have resulted from dams built to confine rivers. As a result of dam construction many once beautiful river valleys have been covered with water and productive farmland in river valleys has been lost. For example, the beautiful Hetch-Hetchy Valley in Yosemite National Park in California has been inundated since the early 1900s by construction of a dam built to provide water and hydroelectric power to San Francisco. Serious proposals are now being considered to remove the dam and restore the valley to its former beauty.

Standing Bodies of Water

Much of the water that humans use comes from standing bodies of water including natural lakes and reservoirs constructed by placing dams on rivers. **Wetlands** are bodies of water shallow enough to support the growth of bottom-rooted plants. **Estuaries** form where fresh river water flows into oceans. They have unique physical, chemical, and biological properties because of the mixing of fresh water and saltwater. Wetlands and estuaries are important breeding grounds for a number of organisms and it is

crucial that they be preserved. Many wetland areas have been drained for agricultural land and an important effort in conservation is their preservation and restoration.

An important characteristic of a lake or reservoir that develops during the summer as a result of the temperature/density behavior of water is **stratification**, shown in connection with water chemistry in Figure 9.4. Exposed to atmospheric oxygen and light, the top layer, the epilimnion, normally has a significant concentration of dissolved oxygen and is **oxic**. Photosynthetic algae

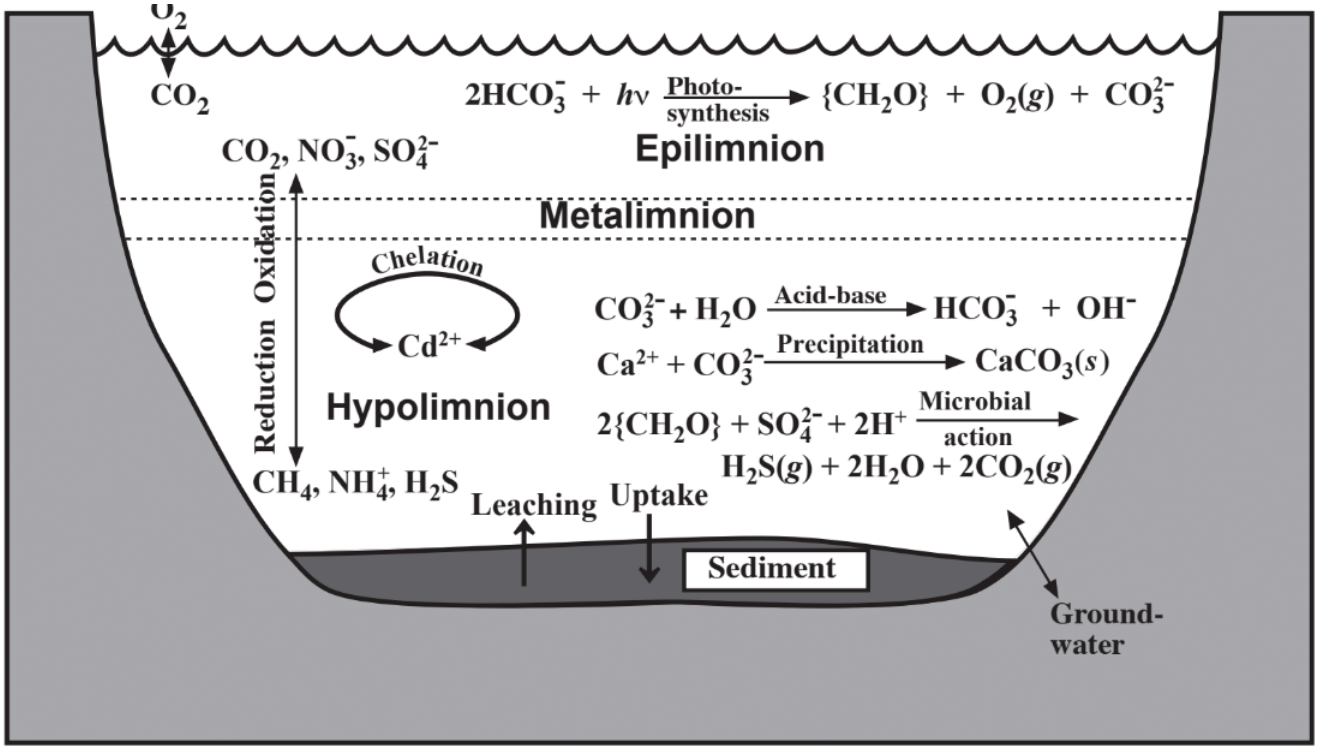


Figure 9.4. Water chemistry as affected by stratification of a body of water in which a warmer, less dense layer of water in the epilimnion resides above a cooler, more dense hypolimnion. This stratification prevents oxygen from penetrating to lower depths where chemically reduced species (CH_4, NH_4^+, H_2S) predominate. In the oxic epilimnion chemically oxidized species (CO_2, NO_3^-, SO_4^{2-}) are present.

thrive in the epilimnion and during daylight produce O2 by photosynthesis. Isolated from atmospheric oxygen, the hypolimnion located in the bottom regions of the body of water becomes **anoxic** as O_2 is consumed by bacteria. Chemically oxidized species including CO_2, NO_3-, and SO_4^{2-} predominate in the epilimnion and chemically reduced species such as CH_4, NH_4^+, and H_2S are found in the hypolimnion. During fall in temperate climates, cooling of the epilimnion causes it to become more dense and to sink resulting in **overturn** of the body of water. This phenomenon tends to stir up sedimentary material and release nutrients from the sediment into the water.

Groundwater

Figure 9.3 illustrates the pathway of groundwater into **aquifers** from precipitation falling on a watershed and infiltrating the aquifer through a recharge zone. The zone of saturation consists of the geospheric rock and soil layers in which all the pores are filled with water, the top of which defines the water table. The fraction of the aquifer formation consisting of pores is the **porosity** and the ability of the formation to allow movement of water is its **permeability**. Generally, high permeability and high porosity are desirable properties of aquifer formations that allow relatively large amounts of water to be withdrawn through water wells drilled into the aquifers. The quality of groundwater is strongly affected by its contact with mineral formations in the geosphere. Infiltration through rock and sand purifies water filtering out microorganisms and suspended solids. Contact with geospheric solids largely determines the chemical composition of water by adding desirable levels of dissolved Ca^{2+} ion and alkalinity and sometimes undesirable solutes such as sodium chloride.

Artesian Wells

Artesianaquifers are those confined by dense clay or shale such that water flows naturally from an **artesianwell** drilled into them. The name comes from the Roman city of Artesum at the site of the French town of Artois known for the free-flowing water from wells dug in the Middle Ages. Artesian wells are often highly prized sources of water. Some modern day bottling companies have leased or purchased sites of artesian wells and advertise their product as artesian water, although it does not differ in quality from pumped well water. Many artesian wells have lost their free-flowing qualities because of depletion of their water resource.

This page titled 2.5.9.2: Occurrence, Availability , and Utilization of Water on Earth is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **9.2: Occurrence, Availability , and Utilization of Water on Earth** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.9.3: Water Chemistry

Some of the chemical and biochemical phenomena that occur in water are illustrated in Figure 9.4.[2] Physically, the body of water shown is stratified during summer months with a warmer, less dense epilimnion floating over a colder, more dense hypolimnion, with little mixing between the two. The epilimnion is in contact with the atmosphere and has a significant content of dissolved oxygen. Therefore, oxidized inorganic species predominate in the epilimnion. The hypolimnion is anoxic because microorganisms consume all the dissolved O_2 in it and it is not in contact with the atmosphere. Reduced inorganic species predominate in the hypolimnion. A major factor in the chemistry of this system is the biochemical photosynthetic production of organic matter represented $\{CH_2O\}$. Organic matter is a biochemical reducing agent and when it sinks into the hypolimnion it is oxidized by microorganism-mediated processes that, for example, reduce NO_3^- and SO_4^{2-} to NH_4^+ and H_2S, respectively. Two important microbially-mediated oxidation-reduction reactions of$\{CH_2O\}$ are reaction with dissolved O_2,

$$(CH_2O) + O_2 \rightarrow CO_2 + H_2O$$

which depletes dissolved oxygen in water making the hypolimnion anoxic and methane fermentation,

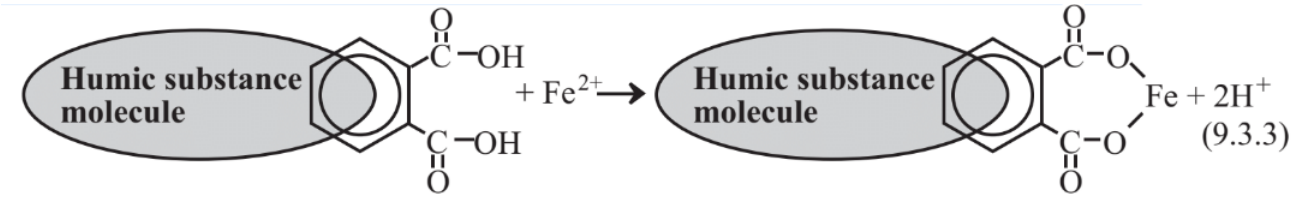
$$2\,(CH_2O) \rightarrow CH_4 + CO_2$$

which produces combustible methane gas. The ability of $\{CH_2O\}$ to react with dissolved O_2 is a measure of the potential of water to become depleted in the oxygen needed by fish and other aquatic organisms and is expressed as **biochemical oxygen demand**, **BOD**, an important water quality parameter.

The photosynthetic biochemical production of biomass results in some important chemical reactions. As illustrated in Figure 9.4, algae use dissolved HCO_3^- ion as a carbon source and in so doing produce carbonate ion, CO_3^{2-}. Two additional reactions of carbonate ion are shown in Figure 9.4. One is its hydrolysis reaction with H_2O molecules back to HCO_3^- with production of OH^- ion. This makes the water basic, an important aquatic acid-base reaction. A second reaction of carbonate is that with dissolved Ca^{2+} to produce solid $CaCO_3$, an important precipitation reaction in water that has been responsible for formation of large deposits of limestone.

All major oxidation-reduction reactions in natural water are carried out by microorganisms acting as catalysts and extracting energy released by the reactions. An example shown in Figure 9.4 is the biochemical reduction of SO_4^{2-} to H_2S. in which sulfate acts as an oxidizing agent to oxidize biomass ($\{CH_2O\}$) in the absence of molecular O_2. Taking place in the hypolimnion and sediments, this reaction is responsible for the foul odor of hydrogen sulfide in some bodies of water and swamps.

Another important phenomenon in natural water and wastewater is the formation of metal chelates in which metal ions are bound in two or more places by organic substances. Humic substances produced by the partial biodegradation of biomass are complex large organic molecules with numerous aromatic ring structures containing oxygen in functional groups including carboxyl (-CO2H) and phenolic hydroxyl (-OH). These groups can lose H+to produce negatively charged groups capable of bonding with metal ions as shown below for the chelation of Fe^{2+} ion:



(9.3.3)

The most important humic substances in water are the lower-molecular-mass fulvic acids. These species tend to chelate Fe^{2+} ion producing a yellow material called *gelbstoffe* (German for “yellow stuff.)” Metal ions bound with fulvic acid are hard to remove from water and, since iron is a very undesirable water impurity, drastic measures such as destruction of the fulvic acid with chlorine may be required to remove the chelated iron.

Phase Interactions in Aquatic Chemistry

Figure 9.4 illustrates the process of exchange of dissolved solutes in water with sediments. Interactions between water and solid, gas, and other liquid phases are very important in aquatic chemistry. Aquatic biochemical processes involve exchange of materials

between water solution and cells of microorganisms. For example, when photosynthesis occurs in water (Figure 9.4), dissolved HCO_3^- ion is transferred into a cell of floating photosynthetic phytoplankton for conversion to biomass. As a product of this reaction gaseous O_2 is released from the cell, some of it dissolving in water and some floating to the top as O_2 bubbles. As noted above, the CO_3^{2-} ion generated as a byproduct of photosynthesis reacts with dissolved Ca^{2+} ion to produce solid $CaCO_3$. The reverse process occurs when bacterially-produced dissolved carbon dioxide reacts with solid calcium carbonate.

$$CO_2\,(aq) + H_2O + CaCO_3\,(s) \rightarrow Ca^{2+}\,(aq) + 2\,HCO_3^-\,(aq)$$

to put calcium ion and bicarbonate ion into solution. Both of these species are important in water. The Ca^{2+} ion is responsible for **water hardness**, named for a tendency to form precipitates with soap anions that are useless for cleaning, and HCO_3^- is **water alkalinity**, the ability to neutralize acids.

A particular kind of phase that interacts with water consists of colloidal particles, which are very small particles of the order of a micrometer in size that are suspended in water. There are three kinds of colloidal particles as shown in Figure 9.5. Many significant materials exist as colloidal particles in water including bacterial and algal cells, clay minerals, soap, and a variety of pollutants. The behavior of colloids is intermediate between that of true solutions and bulk materials such as those in sediments. This behavior is strongly influenced by the very high surface-to-volume ratio of colloids. Aggregation of colloidal particles is an important and often challenging aspect of water treatment, for example in the settling and separation of microorganisms involved in the biological treatment of wastewater.

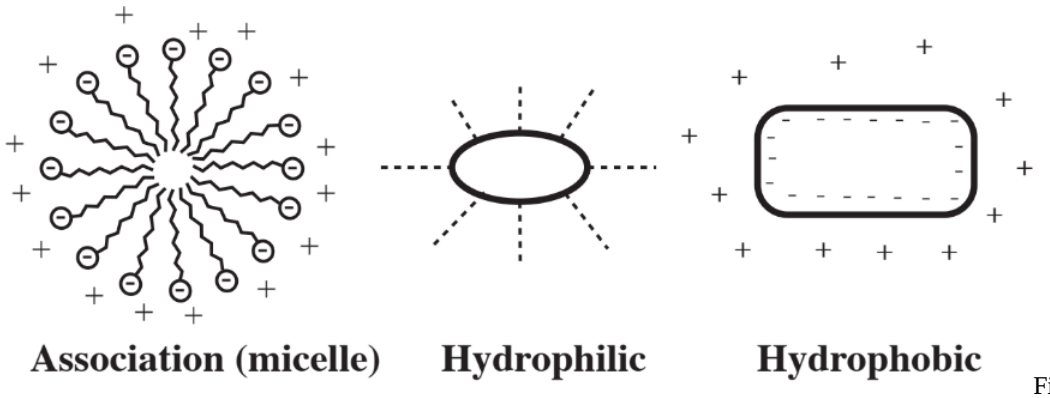


Figure 9.5. Representation of three kinds of colloidal particles in water. Association colloids are composed of micelles in which long hydrocarbon chains on fatty acid anions cluster in the center of the colloid and the anionic functional groups on the ends of the hydrocarbon chains are located around the periphery of the micelle in contact with water. Soap forms association colloids in water. Hydrophilic colloids are materials such as proteins or bacterial cells that are stabilized by their bonding to water. Hydrophobic colloids are repelled by water, but are stabilized by their electrical charges which prevent them from aggregating together.

Acid-Base Phenomena in Water

Natural water almost always contains acids capable of releasing H^+ ion and may contain solutes that can accept H^+ ion and thus act as bases. The most common acidic substance in water is dissolved CO_2, which may enter water from the atmosphere or in higher concentrations as a product of the microbial decay of organic matter. Dissolved carbon dioxide produces H^+ ion by the following reaction with water:

$$CO_2\,(aq) + H_2O \rightleftharpoons H^+ + HCO_3^-$$

The double arrows denote that the reaction is reversible. Since carbon dioxide is a weak acid, the equilibrium of the reaction lies to the left as expressed by the following acid dissociation constant in which $pK_{a1} = -\log K_{a1}$:

$$K_{a1} = \frac{[H^+][HCO_3^-]}{[CO_2]} = 4.45 \times 10^{-7} \quad pK_{a1} = 6.35$$

The HCO_3^- ion can also lose H^+:

$$HCO_3^- \rightleftharpoons H^+ + CO_3^{2-}$$

$$K_{a2} = \frac{[H^+][CO_3{}^{2-}]}{[HCO_3{}^-]} = 4.69 \times 10^{-11} \quad pK_{a2} = 10.33$$

Otherwise pure water in equilibrium with air contains some dissolved CO_2 from the atmosphere, which is 390 parts per million CO_2 by volume. Solubility calculations can be used to show that the dissolved carbon dioxide concentration water, [CO_2(*aq*)], in equilibrium with air is 1.276×10^{-5} M (moles/liter). When each CO_2 reacts with H_2O, one H^+ and one HCO_3^- areproduced. Substitution into the equilibrium constant expression Equation 9.3.6 leads to [H^+] =2.38×10^{-6} M, pH = 5.62, slightly more acidic than neutral pH 7. Therefore, rainwater is naturally slightly acidic. Natural water with a pH less than 5.62 likely contains pollutant strong acid, most commonly H_2SO_4 from acid rain.

Most water that has been in contact with the geosphere contains alkalinity, the ability to react with H^+ ion and neutralize acidity. Generally alkalinity is due to the presence of HCO_3^- ion which undergoes the following reaction with H^+ ion:

$$HCO_3{}^- + H^+ \rightarrow CO_2 + H_2O$$

Alkalinity is normally introduced into water by the reaction of dissolved CO_2 with $CaCO_3$ mineral as shown in Reaction 9.3.4. Because of the presence of alkalinity, most natural waters such as those used to supply municipal water systems are slightly basic with a pH around 8 rather than being slightly acidic like rainwater.

This page titled 2.5.9.3: Water Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **9.3: Water Chemistry** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.9.4: Water Pollution

Water receives a number of different kinds of pollutants and moving water is one of the main pathways of pollutant transport. Water pollutants are of concern for a variety of effects. These include general water quality, toxicity, effects on aquatic organisms, esthetics, eutrophication, and water oxygen levels. Some of the main categories of water pollutants are discussed here.

Several kinds of pollutants are detrimental to water that is a source of water supply. **Taste**, **odor**, and **color** are caused by a variety of impurities. **Sediments** require removal and sometimes special measures are needed to coagulate and settle colloidal suspended matter. Acidity, alkalinity, and salinity at excessive levels are water pollutants.

Some pollutants are picked up by water as a natural consequence of its use in municipal water systems, which take in potable water fit to drink and discharge wastewater. One of the main ones of these consists of sewage including human wastes, food wastes, and other substances that get into water through bathrooms and kitchens. A major component of these water pollutants consists of biodegradable organic matter, abbreviated here as $\{CH_2O\}$ and commonly called **biochemical oxygen demand**, **BOD**, which rapidly depletes receiving water of its dissolved oxygen by the following microbially-mediated process:

$$(CH_2O) + O_2 \rightarrow CO_2 + H_2O$$

Another class of pollutants picked up in municipal water systems and produced in the biodegradation of sewage is composed of **algal nutrients**, primarily nitrogen (NH_4^+, NO_3^-), phosphates ($H_2PO_4^-$, HPO_4^{2-}), and potassium (K^+). Rather than being toxic, these materials cause algae to grow in excess in receiving waters producing too much biomass which degrades as shown in Reaction 9.4.1 and depletes the dissolved oxygen in the water, a process called **eutrophication**. **Detergents** are common pollutants in municipal wastewater and can be harmful because of their content of **surfactants** (basically, the ingredient that lowers water surface tension making it“wetter”) and **builders** added to assist detergent action, formerly a major source of phosphates in wastewater.

A variety of trace **inorganic compounds** can be water pollutants. These include algal nutrients and salts responsible for excess acidity, alkalinity, and salinity mentioned above. Other inorganic pollutants include odorous hydrogen sulfide (H_2S), oxygen-demanding sulfite (SO_3^{2-}),and, in rare cases from mineral processing waste, highly toxic cyanide, CN^-. Trace levels of a variety of elements can be undesirable pollutants. Especially significant of these are toxic **heavy metals** including lead, cadmium, and mercury and metalloid arsenic. **Organically-bound metals** are undesirable. As noted in Section 9.3, iron bound with humic substances causes water color and is particularly difficult to remove. **Methylated forms** of mercury and arsenic (those containing a $-CH_3$ group bound to the element) can be mobilized from sediments and get into water. Dimethylmercury, $Hg(CH_3)_2$, is particularly toxic and undergoes biomagnification through the food chain, accumulating in fish tissue. In addition to the methylated compounds, other **organometallic compounds** made synthetically can be troublesome toxic water pollutants. Until it was banned, tetraethyl lead used as a gasoline additive could get into water and more recently organotin compounds used in biocidal paints to prevent growth of organisms on ship hulls have been toxic to sediment-dwelling organisms.

In some cases **radionuclides** occur as water pollutants. Leakage of radioactive tritium, a form of hydrogen, has gotten into water from reactors. The radionuclide of most concern usually is radium from natural sources and a number of groundwater sources of water have been discontinued because of the presence of radium.

A variety of **organiccompounds** occur as water pollutants. These include industrial chemicals, petroleum products (especially significant in light of the notorious 2010 leak from the Deepwater Horizon oil spill that dumped as much as 5 million gallons of crude oil into the Gulf of Mexico and caused billions of dollars in pollution damage to the Gulf and its shore areas), polychlorinated biphenyls (PCBs, now banned from manufacture but still sometimes encountered, especially in sediments), and dioxin. **Pesticides** are common water pollutants of which **herbicides** such as atrazine are a concern because of their widespread application to land. Some organic water pollutants call for particular attention because of their potential biological activity and toxicity. **Carcinogens** are an obvious case. **Pharmaceuticals** and their metabolites have emerged as a concern and are currently one of the “hottest” areas of water pollution research. A wide variety of these get into wastewater and sometimes into water supplies, though in minuscule quantities.

This page titled 2.5.9.4: Water Pollution is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **9.4: Water Pollution** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.9.5: Greening of Water - Purification Before Use

From the air the San Marcos/ Cazones River near the city of Poza Rica, Mexico, has a beautiful green color. But this river's water is not what normally would be chosen as a water source (although by necessity it is so used), the reason being that its appearance denotes a major pollution problem because its green color indicates excessive growth of plants and algae (eutrophication) in the river, nourished by untreated sewage and other wastes discharged into it. This chapter discusses keeping water green, not in color but in terms of water quality.

For most uses water requires treatment. The processes used to treat water are generally similar for municipal, commercial, or industrial uses. This section addresses the major treatment processes applied to water supplies.

Municipal Water

The first steps in water treatment are usually physical, blowing air through it (aeration) and sedimentation to allow solids to settle from the water. Another benefit of water aeration is the oxidation of soluble Fe^{2+}, a very undesirable pollutant that causes staining of bathroom fixtures and clothing, to solid $Fe(OH)_3$, which settles from the water. A benefit of the formation of this gelatinous solid is that it acts as a **coagulant** that binds with colloidal impurities entrained in the water and causes them to settle. Another physical process to which municipal water is usually subjected is **filtration** to filter out solids. Filtration is usually over sand filters, but activated carbon, a form of carbon treated at high temperatures with steam or carbon dioxide to create pores and give it an enormous internal surface area, may be used as a filtration medium to remove organic substances.

Some water sources have too much dissolved Ca^{2+} (water hardness) that forms precipitates with soap and deposits solid $CaCO_3$ in pipes. Normally the dissolved calcium is present as calcium bicarbonate in which the anion present with Ca^{2+} is bicarbonate, HCO_3^-. Curiously, the method for removing this calcium is to add more calcium in the form of lime, $Ca(OH)_2$, a basic substance that converts the HCO_3^- to CO_3^{2-} and removes the dissolved calcium as solid $CaCO_3$ according to the following reaction:

$$Ca^{2+}(aq) + 2\,HCO_3{}^-(aq) + Ca^{2+}(aq) + 2\,OH^-(aq) \rightarrow 2\,CaCO_3(s) + 2\,H_2O$$

The final step in water purification is **disinfection**, commonly by the addition of elemental chlorine gas, which reacts with water,

$$Cl_2 + H_2O \rightarrow HOCl + H^+ + Cl^-$$

to produce disinfectant HOCl that kills virus and bacteria in water. Hypochlorite salts of HOCl including NaOCl and $Ca(OCl)_2$ can also act as disinfectants. Chlorine and hypochlorites can react with low concentrations of added ammonia to produce chloramines such as NH_2Cl that remain in the water distribution system and maintain residual disinfection In the presence of residual organic matter such as humic substances, elemental chlorine may form potentially toxic chloroform, $HCCl_3$, and a class of related substances commonly called trihalomethane compounds. To prevent these compounds from forming, a common practice is the use of chlorine dioxide, ClO_2, which does not produce trihalomethanes. Chlorine dioxide is too dangerously reactive to move and too unstable to store so it is made on site by the reaction of sodium chlorite with elemental chlorine:

$$2\,NaClO_2(s) + Cl_2(g) \rightarrow 2\,ClO_2(g) + 2\,NaCl$$

Therefore, potentially dangerous chlorine dioxide is prepared only when needed, in the quantities needed, where needed, which is in keeping with the best practice of green chemistry and technology.

Green Ozone

A greener alternative to chlorine-based water disinfectants in many respects is ozone, O_3. Pumped into water, this form of oxygen kills pathogens without producing the undesirable byproducts made by chlorine and it is actually more effective than chlorine in killing viruses. Ozone is produced from oxygen in air by a high-voltage electrical discharge through dried air as illustrated in Figure 9.6. The lifetime of ozone in water is short, so a small amount of chlorine must usually be added to ozonated water to maintain disinfection in the water distribution system.

Disinfection of water by ozonation is a virtually ideal example of green chemical practice. The only raw material is universally available air, which is free. Ozone is produced only where it is needed as it is needed, without byproducts. The ozone does not persist in water, decomposing to elemental oxygen. There is very little likelihood of producing harmful disinfection byproducts with ozone.

Water for Commercial and Industrial Use

A wide range of water quality is required for water destined for commercial and industrial use and for economic reasons the water is treated only as needed for its intended application. Water used for cooling generally requires little treatment, the main requirement being that it is wet. At the other end of the scale water used in the semiconductor industry has to be hyperpure. Pathogens obviously must be removed for water employed in food processing and corrosive and scale-forming contaminants must be removed from boiler feedwater.

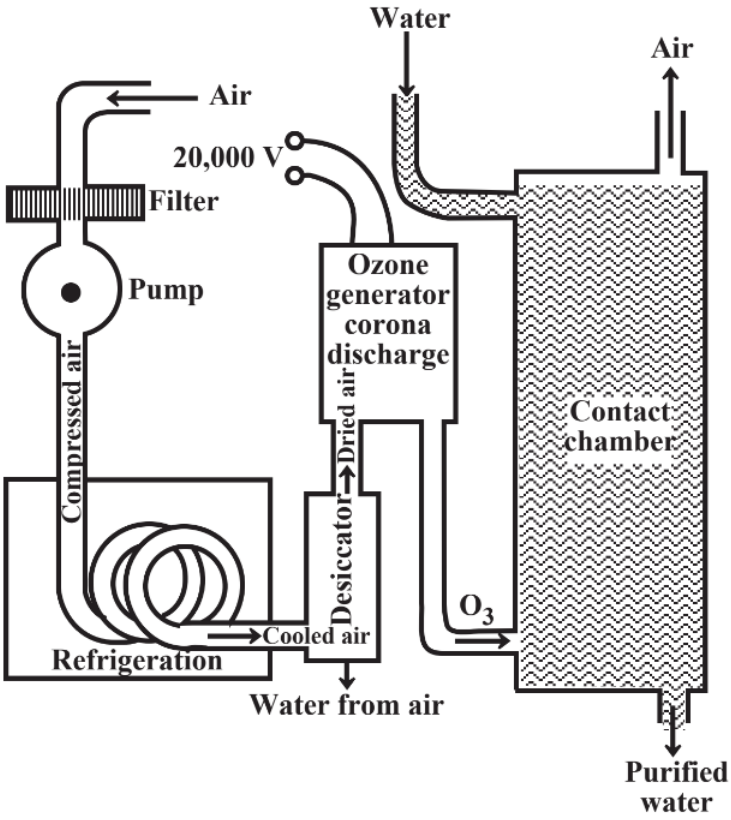


Figure 9.6. A water disinfection system based upon ozone produced by an electrical discharge through chilled, dried air followed by pumping the ozonated air through water.

The same treatment processes described above for municipal water are applied to water destined for commercial and industrial use. Some other treatment steps commonly applied to commercial and industrial water supplies include the following:

- Addition of precipitants such as Na_3PO_4 to remove Ca^{2+} and prevent formation of $CaCO_3$
- Addition of dispersants to prevent scale formation
- Adjustment of pH by addition of acid or base
- Disinfection to remove pathogens and addition of biocides to prevent microbial growths
- Addition of coagulants followed by filtration to remove suspended colloidal material
- Treatment with activated carbon to remove organics
- Deionization to remove salts
- Reverse osmosis to remove salts

For economic and conservation reasons significant amounts of water used for commercial and industrial applications are recycled as discussed in Section 9.7. Such water is often subjected to **sequential use** for applications that require successively lower quality, the last use before discharge requiring the lowest quality of water. In some cases water leaving a facility may be applied to grass or golf courses or used for irrigation.

This page titled 2.5.9.5: Greening of Water - Purification Before Use is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **9.5: Greening of Water - Purification Before Use** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.9.6: Wastewater Treatment

Although there are many kinds of wastewater depending upon where the water has been used, the most common wastewater treatment is applied to sewage consisting of water that has been through a municipal water distribution system. The most common contaminant of municipal wastewater consists of biodegradable organic matter, abbreviated $\{CH_2O\}$, usually biological material from human wastes or biomass flushed down the drain by garbage disposal grinders. When this biodegradable matter gets into the receiving waters where sewage effluent ultimately ends up, its biodegradation consumes dissolved oxygen as shown in Reaction 9.3.1. As noted in Section 9.3, biochemically degradable matter in water is said to exert a biochemical oxygen demand, BOD. The main objective of municipal wastewater treatment is reduction of BOD.

There are three main categories of sewage treatment: (1)primary treatment to remove solid objects, grit, and grease, (2) secondary treatment to reduce BOD, and (3) tertiary treatment to further refine the quality of the effluent water. Primary treatment is essentially self-explanatory and tertiary (advanced) treatment is addressed in more detail in Section 9.6 which discusses treatment of wastewater for recycling. The following is a brief discussion of secondary wastewater treatment.

Normally biological treatment with microorganisms is employed to remove BOD in secondary wastewater treatment. In fixed film bioreactors this is accomplished with a film of bacteria and protozoa immobilized on a support so that the microorganisms are alternatively exposed to wastewater and to air. One way in which this is done is with a trickling filter consisting of rocks coated with microorganisms onto which wastewater is drained from a rotating pipe with holes along the bottom edge. A second means is with microorganism-coated disks that rotate on a shaft so that half of each disk is immersed in water and the other half exposed to air at any given time.

The most widely used means of secondary wastewater treatment is the **activatedsludge process** shown in Figure 9.7. In this process, biomass is degraded in a tank containing a relatively large mass of microorganisms kept suspended by air pumped into the bottom of the tank. This air serves as a continually renewed source of dissolved molecular oxygen in the water. In the aeration tank biodegradable organic material is acted upon by the microorganisms to produce biomass and carbon dioxide. As part of this process organically bound nitrogen, sulfur, and phosphorus are largely converted to simple inorganic species, such as NH_4^+ from organic nitrogen. The treated wastewater exits the tank after an appropriate residence time and goes to a settling basin in which the suspended mass consisting largely of bacterial and protozoal biomass called sewage sludge(now often given the less offensive term of biosolids) settles and the treated water is discharged. The settled sewage sludge is then pumped back to the front of the aeration tank to provide a constant high population of biodegrading organisms in the tank. As the process progresses, the mass of sewage sludge increases and the excess is transferred to an anoxic digester where it remains for some time and generates methane (CH_4) by anoxic fermentation in the absence of oxygen. The methane is used as a fuel and can be employed to run engines that power the plant. Spent sludge accumulates in the anoxic digester and requires disposal. The “greenest” means of disposal is to spread it onto farmland as fertilizer. Often it is incinerated, which requires significant amounts of supplemental fuel because the sludge has such a high water content.

This page titled 2.5.9.6: Wastewater Treatment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **9.6: Wastewater Treatment** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.9.7: Advanced Wastewater Treatment and Recycling

Advanced or tertiary water treatment is applied to the effluent from secondary wastewater treatment facilities. Such treatment is becoming more common in part because water from secondary treatment often contains sufficient BOD to cause problems in receiving waters and contains excessive levels of algal nutrients, especially phosphate, that results in excess algae growth and eutrophication, a condition in which algal biomass accumulates in excess, then decays and depletes water of dissolved oxygen. The other reason for advanced wastewater treatment is that wastewater can be a readily available water source that is potentially easier to bring up to drinking water standards compared to seawater or brackish groundwater. A major barrier to complete water reuse is public opposition to the idea of using water that once was sewage although that is exactly what is done in cases where cities draw their water supplies from rivers downstream from where other cities have discharged treated wastewater.

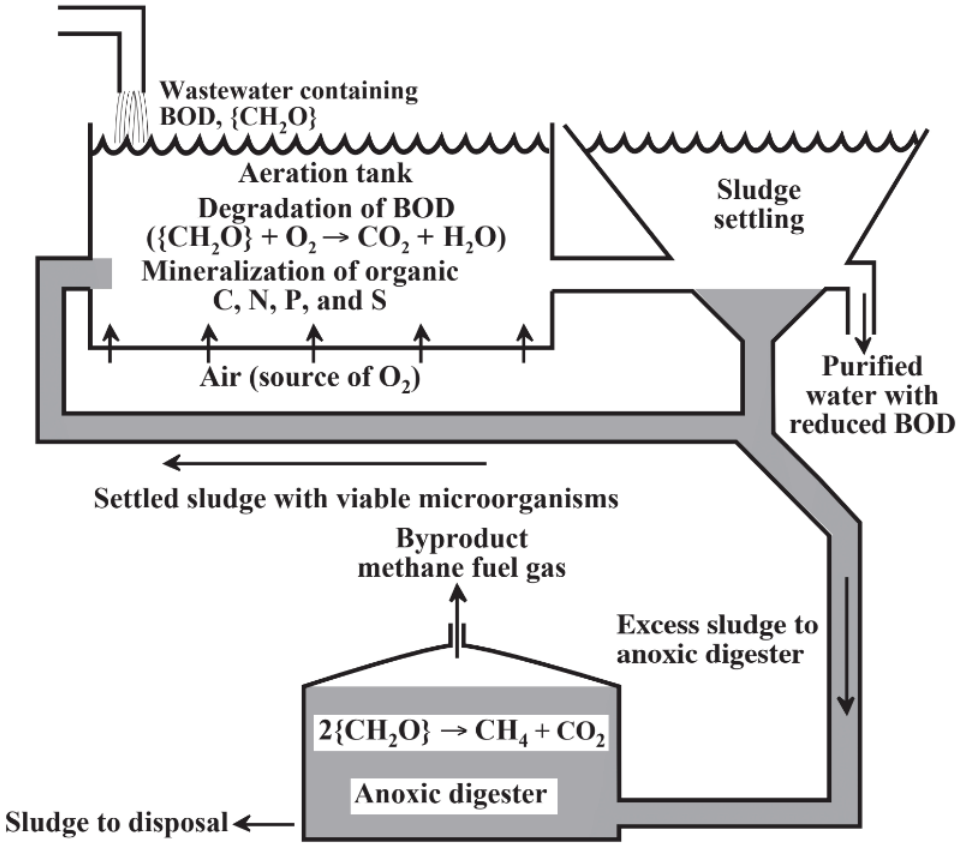


Figure 9.7. An activated sludge wastewater treatment facility

Most of the water treatment processes discussed so far in this chapter can be used for advanced wastewater treatment. The two major contaminants that need to be removed from such water are organics including BOD and salts including algal nutrient ions containing phosphorus or nitrogen. Also of concern are pathogens including virus and waste pharmaceuticals and their byproducts. Much of the BOD is in the form of residual suspended solids that can be removed by filtration processes. Filtration over beds of sand has long been used to purify water. More recently filtration with membranes that remove extremely small particles has become much more popular. Nanofiltration through membranes with pores 0.1–5 nanometers (nm) in size can even remove viruses from water, a step that is very important for recycling water from sewage sources. Filters with even smaller pores are used for **reverseosmosis** treatment of water (Figure 9.8) which removes salt ions including algal nutrients with a semipermeable membrane that attracts pure water but rejects ions. Reverse osmosis produces a concentrated brine product that requires disposal.

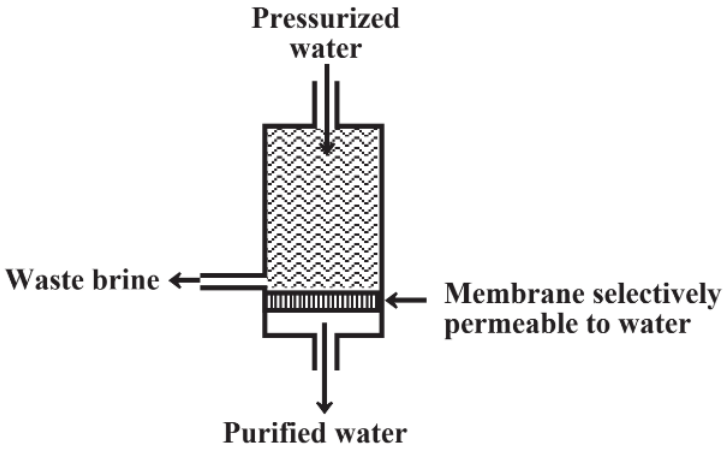


Figure 9.8. A reverse osmosis system for the removal of ions and other impurities from water. Membrane filtration processes that employ filters with larger pores are used to remove a variety of contaminants from water.

Figure 9.9 shows a system for reuse of water capable of producing drinking water from sewage. It illustrates the major kinds of treatment processes that can be employed and the applications of the water product. Taking effluent from a secondary sewage treatment plant, this system takes advantage of nature's inherent ability to purify water by first placing the water in a constructed wetlands. Plants and algae grow profusely in the nutrient-rich wastewater effluent, and their biomass can be harvested for energy production. Water from the wetlands can percolate into underground aquifers or can even be pumped into aquifers where contact with mineral formations further purifies the water. Water leaving the wetlands can be used without additional treatment for irrigation or other non-potable applications, such as cooling water. For household or other uses requiring a higher purity product, the water can be treated with activated carbon to remove organics and subjected to membrane filtration processes to remove small particulate contaminants and some kinds of large-molecule solutes and pathogens, including virus. To remove dissolved salts, reverse osmosis may be employed. Ultraviolet irradiation of the water can be used to destroy pharmaceuticals and their metabolites. The final water product can then be introduced into a municipal water system.

The 275 million liter/day Groundwater Replenishment System of the Orange County Water District in southern California, which cost $481 million and went into operation in late 2007, is the world's largest water plant designed to produce potable quality water from treated sewage effluent. The water from this system is not currently used for municipal water supply but is pumped into underground aquifers to serve as a future source of drinking water and to prevent infiltration of saline ocean water into the aquifers. The operation of this plant is expected to serve as a model for similar plants around the world.

This page titled 2.5.9.7: Advanced Wastewater Treatment and Recycling is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **9.7: Advanced Wastewater Treatment and Recycling** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.9.8: The Many Uses of Water

Beyond its well known domestic uses and for irrigation, water has many applications. A renewable, sustainable green resource, next to air water is the cheapest and most universally available raw material and, because it is totally renewable, it is the greenest of raw materials. Here are addressed just a few of the huge number of water-based technologies.

Many of water's uses come from its unique physical properties, especially as they relate to water's extraordinarily high capacity to absorb and release heat energy. At 100°C,a very high 40.7kilojoules/mole (kJ/mol) is required to change liquid water to steam and a correspondingly large amount of heat is released when steam condenses to liquid water, the highest heat of vaporization of any common substance. Advantage is taken of this characteristic by vaporizing water in a boiler, moving it as steam, and condensing it back to liquid, a very effective means of moving heat energy from one place to another, such as for heating multiple buildings on a university campus, for example. Some European cities use such a **districtheating** system to heat residential houses and apartments, often with boilers fueled in part by municipal refuse. The high heat of fusion for ice enables the use of solid water to absorb heat.

Steam is a very effective **workingfluid** to convert heat energy to mechanical energy. This used to be done with piston engines including those on steam locomotives, but is now performed by steam turbines. Steam is produced in **water-tube** boilers in which liquid water is evaporated inside 2-8 cm diameter tubes around which hot combustion gases are circulated. Water drums for water storage and steam drums for steam storage are isolated from the fire to prevent an explosion in case they rupture, which would happen if their contents got into the firebox.

An interesting application of steam which greatly facilitates the sustainable use of wood is **steam explosion pulping** of wood in which wood is heated with pressurized steam at 180°C to 210°C. This step is followed by very rapid decompression such that the wood literally explodes releasing wood fibers that can be used in numerous applications including making paper pulp and fiber board. The cellulose in the fibers can be treated to produce sugars that can be fermented to produce alcohol.

An important consideration in producing steam for any purpose is the prevention of scaling, fouling, and corrosion, which generally requires very pure and carefully treated water. The best source of such water is the steam itself, which is condensed for **makeupwater** to return to the boiler. The feed water to a boiler is treated to remove contaminants that can cause solid buildup in the boiler tubes. Phosphate salts are added to remove dissolved calcium and magnesium (hardness) by reactions such as the following:

$$5\,Ca^{2+}(aq) + 3\,PO_4{}^{3-}(aq) + OH^-(aq) \longrightarrow Ca_5OH(PO_4)_3(s)$$

The removal of silicon from boiler feedwater is required because it can carry over with steam and cause damaging deposits of SiO_2 on turbine blades. Both dissolved carbon dioxide and oxygen are corrosive. Hydrazine can be used to remove traces of residual oxygen by the following reaction:

$$N_2H_4 + O_2 \rightarrow 2\,H_2o + N_2$$

Anticorrosive agents, such as cyclohexylamine are added to boiler feedwater.

Videos of floodwaters uprooting structures and carrying them away attest to water's ability to transmit mechanical energy. Advantage is taken of this capability by using jets of pressurized water to harvest sand and gravel from river beds and to mine and process some kinds of minerals.

The greatest use of moving water's mechanical energy is the application of **waterpower**. Employed for many centuries as a renewable energy source, waterpower is the oldest source of non-animal energy harnessed, for example, by waterwheels powering grain-grinding applications or sawmills. Energy may be harvested from water flowing downhill from a dammed stream and from rising and falling ocean tides. The greatest current use of moving water is for **hydroelectric power** to run electricity generators. The U.S. now has about 100,000 megawatts of hydroelectric capacity (a large fossil-fueled or nuclear power plant typically has a capacity of 1000 MW). This capacity could be doubled by using all available sites in the U.S., including Alaska for hydroelectric power, although this will not happen for economic and environmental reasons. A useful adaptation of hydroelectric power is **pumpedstorage** in which turbines are reversed to pump water to containment structures at higher elevations during periods of low electricity demand and the stored water is used to run turbines attached to electrical generators when demand is high.

Water power affords economic and environmental advantages including a free source of “fuel” and the lack of emissions or ash, making waterpower one of the “greenest” sources of energy. Problems may arise under extreme drought conditions when water is simply unavailable.

Unfortunately, the vast reservoirs required for most waterpower developments destroy free-flowing rivers and are detrimental to fish migration, such as that required for salmon reproduction. Most potentially remaining available sites are in remote regions from which the transfer of electricity to population centers requires massive power lines, which also present environmental problems. Because of this there is now a trend toward dismantling dams to restore normal stream flow, the esthetics of river valleys and gorges, and wildlife habitat. (Sediments accumulated in reservoirs, some of which contain hazardous material, can be released when dams are removed and cause problems downstream.) New technologies have also been developed that utilize the power of moving water in river beds without the need for dams.

Water as a Solvent

Water is an extraordinarily good solvent for a variety of materials and its solvent properties can be enhanced with suitable additives, especially surfactants that reduce its surface tension making it a powerful cleaning agent. Mixed with suspended lubricants, water is very useful as a lubricant and cooling agent, such as in metal-stamping operations. Substitution of appropriately treated water for organic solvents, which generally come from nonrenewable petroleum and are expensive, has extended water’s use for washing small parts, such as electronic constituents. As a solvent for chemical reactants, water serves as a medium for many chemical synthesis and processing applications. Substances can be purified by dissolving them in water, then evaporating some of the water off to leave a purified form of the substance; some salts are purified by this means.

Water is useful for its chemical properties and is a chemical ingredient for a number of industrial chemical reactions. It is required for the hardening of Portland Cement to make concrete. It can be used as the reagent in treating some kinds of hazardous wastes by hydrolysis. Water serves as a source of elemental hydrogen used as a fuel in fuel cells and as a raw material in making some chemicals, such as ammonia. Elemental hydrogen is generated along with oxygen gas when a direct electrical current is passed through water that has been made conducting by the addition of a salt. Hydrogen is produced at the negatively charged cathode when electrons (*e*-) are added to molecules of water,

$$2\,H_2O + 2\,e^- \rightarrow H_2\,(g) + 2\,OH^-$$

and oxygen is generated at the positively charged anode by the removal of electrons from water:

$$2\,H_2O \rightarrow O_2\,(g) + 4\,H^+ + 4\,e^-$$

The net reaction is simply the following:

$$2\,H_2O \rightarrow 2\,H_2\,(g) + O_2\,(g)$$

The process is certainly an example of green chemistry because the only reagent is water and the only byproduct is oxygen gas, which has a number of uses or can be released harmlessly to the atmosphere. In Iceland, hydrogen gas made by the electrolysis of water generated from abundant hydroelectric and geothermal sources of electricity is used to fuel automobiles.

This page titled 2.5.9.8: The Many Uses of Water is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **9.8: The Many Uses of Water** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.9.9: Hot Water - Pressurized Subcritical Water

As described in Section 9.10, water heated above its critical temperature of 374.4°C and kept under very high pressure reaches a special state called supercritical fluid. However, at temperatures above its normal boiling temperature of 100°C but below those required to reach the supercritical state, the properties of liquid water change markedly. Whereas the dielectric constant of liquid water at 25°C is 78, that of pressurized subcritical water at 205°C is 33, the same as methanol at 25°C. At temperatures between 100°C and 205°C, water behaves like a mixture of water and methanol and becomes a much more effective solvent for organic substances. Such superheated water has a number of uses such as extracting organic materials from plant biomass, extracting pollutant organic materials from contaminated soil and sediments, and as a reaction medium for organic chemical reactions. The decrease in dielectric constant of water and the increase in solubility of substances at higher temperatures means that the solubilities of some substances in superheated water may be orders of magnitude higher than in water at 25°C.

A potentially useful application of pressurized hot water is its application as a medium along with pressurized O_2 in which organic substances are oxidized. This procedure has been used to make oxygenated products from organics as refractory as coal. It has also been used as a means of destroying oxidizable pollutants in water. For example, the organic matter in sewage can be removed by reaction with O_2 in superheated water.

This page titled 2.5.9.9: Hot Water - Pressurized Subcritical Water is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **9.9: Hot Water - Pressurized Subcritical Water** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.9.10: Supercritical Water

At temperatures exceeding 374.4°C and an extremely high pressure of at least 217.7 atm (217.7×normal atmospheric pressure) the distinction between liquid water and vapor vanishes and water enters a special physical state called a **supercritical fluid**. As a supercritical fluid, water no longer has the abundance of hydrogen bonds that give liquid water its special properties and that make it such a good solvent for ionic materials; instead, it becomes a good solvent for organic materials. The properties of supercritical water can be changed over a wide range by varying temperature and pressure and conditions can be attained in which the water is completely miscible with nonpolar compounds while retaining its ability to solubilize ions and polar species. Varying the pressure on supercritical water can change its density continuously from gaslike to liquid-like with a corresponding variation in the properties of the fluid. Although the conditions required to maintain water as a supercritical fluid are very severe and require special equipment, its ability to substitute for organic solvents make supercritical water very useful for a number of purposes, such as a medium for organic synthesis reactions. Even when compressed to a density equal to that of liquid water, supercritical water has a low viscosity. This is important for chemical processes because it increases mass transfer and rates of diffusion in diffusion-controlled processes.

Supercritical water has enormous potential as a solvent medium and catalytic material for chemical reactions. Its high temperature is conducive to pyrolysis, it promotes the hydrolysis of compounds in which molecules are cleaved with addition of H_2O, and in the presence of O_2 and other oxidants it is a strongly oxidizing medium. Careful tuning of temperature, pressure, and oxidant levels enables supercritical water to facilitate partial oxidation of organic materials such asthe partial oxidation of methane to produce methanol:

$$CH_4 + \tfrac{1}{2}\,O_2 \rightarrow CH_3OH$$

Whereas methane must be moved by pipeline or as an extremely cold cryogenic liquid, which is hard to do from some of the remote regions where natural gas is found, methane converted to liquid methanol can be shipped in large tankers and used as a motor fuel or in fuel cells.

Oxygen, like other common inorganic gases, is completely miscible with supercritical water under extreme pressures, which increases its capability to act as an oxidant in chemical synthesis and waste destruction. Supercritical water is an excellent medium in which to use dissolved molecular oxygen to oxidize organic wastes including chlorinated compounds to carbon dioxide, water, and inorganic halides (chloride ion). The process is facilitated by the ability of supercritical water under certain conditions to act as a good solvent for organic wastes such as polychlorinated biphenyls (PCBs).

Under the extreme pressure and temperature conditions of around 30 km in the geosphere water is likely to be supercritical. In the presence of supercritical water at these depths, minerals may behave much differently compared to their behaviors under normal conditions, especially with respect to their dissolution and precipitation behavior. There is also evidence to suggest that chemical processes may form methane under these conditions leading to a non-biological pathway for hydrocarbon formation.

This page titled 2.5.9.10: Supercritical Water is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **9.10: Supercritical Water** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.9.11: Literature Cited

1. Manahan, Stanley E., *Environmental Science and Technology: A Sustainable Approach to Green Science and Technology*, 2nd ed., Taylor & Francis/CRC Press, Boca Raton, FL, 2006.

2. Manahan, Stanley E., *Water Chemistry: Green Science and Technology of Nature's Most Renewable Resource*, Taylor & Francis/CRC Press, Boca Raton, FL, 2010.

2.5.9.11: Literature Cited is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Literature Cited** is licensed CC BY-NC-SA 4.0.

2.5.9.12: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. Look up proposals to restore the Hetch-Hetchy Valley in Yosemite National Park to its former state. How might this affect water supply to parts of California? What might be some benefits of restoration of this valley to its former state?

2. Look up and explain the significance of the name Mulholland in relationship to water. How did Mulholland affect history?

3. Paradoxically, pollution by a strong acid such as HCl of groundwater in contact with limestone ($CaCO_3$) can lead to an increase in the alkalinity of the water. Using chemical reactions, explain how this may occur.

4. Tests can be performed on water that show the presence of biochemical demand (BOD) and other tests that chemically oxidize organic matter to produce CO_2 can show total organic carbon (TOC). Applied to a particular sample of water, these two tests showed relatively high TOC and relatively low BOD. What does this say about the nature of the organic pollutants in the water?

5. Agricultural fertilizer normally adds nitrogen, phosphorus, and potassium to soil. Explain how fertilizer runoff into a body of water can lead eventually to increased biochemical oxygen demand pollution.

6. The development of the flameless atomic absorption analysis for mercury enabled very sensitive tests for this element around 1970 and showed suprisingly high levels of this toxic element in some fish samples. The inorganic chemistry of mercury suggests that mercury compounds should precipitate and settle into sediments where they are unavailable to fish. What, then was the explanation for the high mercury levels found in some fish around 1970?

7. Phosphate in the form of $H_2PO_4^-$ and HPO_4^{2-} ions is the substance usually removed from secondary sewage effluent to prevent excessive algal growth and eutrophication in receiving waters. Of several possible algal nutrients, why is phosphate chosen? Show with a chemical reaction the most common means of removal.

8. Membrane filtration processes can be very effective in removing residual BOD from secondary wastewater effluent. What does this suggest regarding the nature of contaminants responsible for the BOD?

9. By doing some search on the internet, gather information regarding the use of wastewater for irrigation. Is this a practice that is used and if so where does it usually take place? What are some of the benefits? What are some of the risks?

10. In Section 9.2 information is given about the total amounts of water on Earth and how much of it is in the solid form as ice, snowpack, and glaciers. Information is available on the internet regarding the heat of fusion of ice and the rate at which energy reaches Earth from the sun. From this information, estimate the length of time required to melt all of Earth's solid water if all the solar flux received by Earth could go to that purpose.

11. Water is used for both its special solvent properties and its ability to absorb, transfer, and release heat energy. Explain on the basis of the characteristics of the water molecule how these two uses are related.

12. How far back into history does the use of waterpower go? Which civilizations were the first to use it? Explain why during the mid-1800s waterpower development slowed, only to start growing rapidly around the late 1800s and early 1900s.

13. What is Plaster of Paris? Show with a chemical reaction how water is employed as a chemical reagent in making objects from Plaster of Paris

2.5.9.12: Questions and Problems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** is licensed CC BY-NC-SA 4.0.

2.5.9.13: Supplementary References

Amjad, Sahid, *The Science and Technology of Industrial Water Treatment*, Taylor & Francis/CRCPress, Boca Raton, FL, 2010.

Benjamin, Mark M., *Water Chemistry*, McGraw-Hill, New York, 2002.

Berk, Zeki, *Water Science for Food, Health, Agriculture and Environment*, Technomic Publishing, Lancaster, PA, 2001.

Brezonik, Patrick L., and William A. Arnold,*Water Chemistry: An Introduction to the Chemistry of Natural and Engineered Aquatic System*, Oxford University Press, Oxford, UK, 2011.

Dodds, Walter K., and Matt R. Whiles, *Freshwater Ecology: Concepts and Environmental Applications of Limnology*, 2nd ed., Academic Press, Boston, 2010.

Dodson, Stanley, I., Introduction to *Limnology*, McGraw-Hill, New York, 2005.

Essington, Michael E., *Soil and Water Chemistry: An Integrative Approach*, Taylor & Francis/CRC Press, 2004.

Kalff, Jacob, *Limnology: Inland Water Ecosystems*, Prentice Hall, Upper Saddle River, NJ, 2002.

Kumar, Arvind, Ed., *Aquatic Ecosystems*, A.P.H. Publishing Corp., New Delhi, 2003.

NamiesÌnik, Jacek, and Piotr Szefer, Eds., *Analytical Measurements in Aquatic Environments*, Taylor & Francis/CRC Press, London, 2010.

Nollet, Leo M. L., *Handbook of Water Analysis*, 2nd ed., Taylor & Francis, CRC Press, 2008.

Salina, Irena, Ed., *Written in Water: Messages of Hope for Earth's Most Precious Resource*, National Geographic, Washington, DC 2010.

Spellman, Frank R., *The Science of Water: Concepts and Applications*, 2nd ed., Taylor & Francis/CRC Press, Boca Raton, FL, 2007.

Sullivan, Patrick, Franklin J. Agardy, and James J. J. Clark, *The Environmental Science of Drinking Water*, Butterworth-Heineman, Burlington, MA, 2005.

Trimble, Stanley W., *Encyclopedia of Water Science*, 2nd ed., Taylor & Francis, CRC Press, 2008.

Water Environment Federation, *Industrial Wastewater Management, Treatment, and Disposal*, 3rd ed., McGraw-Hill Professional, New York, 2008.

Welch, E. B., and Jean Jacoby, *Pollutant Effects in Fresh Waters: Applied Limnology*, Taylor &Francis, London, 2007.

2.5.9.13: Supplementary References is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Supplementary References** is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

2.5.10: Blue Skies for a Green Environment

“We are absolutely dependent upon the atmosphere for materials essential to life and for protection from the hostile environment of outer space. This marvelous canopy over us is really very thin and fragile. If Earth were the size of a classroom geography globe, virtually all of the mass of the atmosphere would be contained in a layer about the thickness of the varnish on the globe!

2.5.10.1: More Than Just Air to Breathe

2.5.10.2: The Gas Laws

2.5.10.3: The Protective Atmosphere

2.5.10.4: Atmospheric Chemistry and Photochemical Reactions

2.5.10.5: The Enormous Importance of Climate

2.5.10.6: Atmospheric Particle Pollutants

2.5.10.7: Pollutant Gaseous Oxides

2.5.10.8: Acid Rain

2.5.10.9: Miscellaneous Gases in the Atmosphere

2.5.10.10: Photochemical Smog

2.5.10.11: Natural Capital of the Atmosphere

2.5.10.12: Energy and Mass Transfer in the Atmosphere

2.5.10.13: Questions and Problems

2.5.10.14: Supplementary References

This page titled 2.5.10: Blue Skies for a Green Environment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.10.1: More Than Just Air to Breathe

A Sea of Gas

We live and breathe in the **atmosphere**, a sea of gas consisting primarily of elemental O_2 and N_2. The fundamental properties of gases determine the properties of the atmosphere. Recall that gases consist of molecules and (in the case of noble gases) atoms with large amounts of space between them. The gas molecules are in constant, rapid motion, which causes gases to exert pressure. The motion of gas molecules becomes more rapid with increasing temperature. This motion also means that gas molecules move by a process called **diffusion**. The relationships among the amount of gas in moles, its volume, temperature, and pressure can be calculated by the gas laws discussed in Section 10.2.

Whereas seawater in the ocean has a well-defined volume and a distinct surface, the same cannot be said for the mass of gases comprising the atmosphere. Although most of the atmosphere is within a few kilometers of Earth's surface, there is no distinct point at higher altitude where the atmosphere ends. Instead, air becomes progressively thinner with increasing altitude. This is noticeable to humans who have traveled to higher altitudes on mountains where the thinner air makes breathing more difficult. Indeed, climbers who scale the highest mountain peaks commonly carry oxygen to aid breathing.

Atmospheric Composition

What is air? At our level, it is a mixture of gases of uniform composition, except for water vapor, which composes 1-3% of the atmosphere by volume, and some of the trace gases, such as pollutant sulfur dioxide. On a dry basis, air is 78.1% (by volume) nitrogen, 21.0% oxygen, 0.9% argon, and 0.04% carbon dioxide. Normally, air is 1–3% water vapor by volume. Trace gases at levels below 0.002% in air include ammonia, carbon monoxide, helium, hydrogen, krypton, methane, neon, nitrogen dioxide, nitrous oxide, ozone, sulfur dioxide, and xenon.

By a wide margin, oxygen and nitrogen are the most abundant gases in the atmosphere. Because of the extremely high stability and low reactivity of the N_2 molecule, the chemistry of atmospheric elemental nitrogen is singularly unexciting, although nitrogen molecules are the most common "third bodies" that absorb excess energy from atmospheric chemical reactions, preventing the products of addition reactions in the atmosphere from falling apart. Oxides of nitrogen actively participate in atmospheric chemical reactions. Elemental nitrogen is an important commercial gas extracted from the atmosphere by nitrogen-fixing bacteria and in the industrial synthesis of ammonia.

Oxygen is a reactive species in the atmosphere that reacts to produce oxidation products from oxidizable gases in the atmosphere. Two such species that are particularly important are sulfur dioxide gas, SO_2, and pollutant hydrocarbons. Molecular O_2 does not react with these substances directly but only indirectly through the action of reactive intermediates, especially **hydroxyl radical**, HO•.

A crucially important atmospheric chemical phenomenon involving oxygen is the formation of stratospheric ozone, O_3. The formation of this gas in the stratosphere is discussed in Section2.13 and shown by Reactions 2.13.1 and 2.13.2.

Oxygen in the atmosphere is consumed in the burning of hydrocarbons and other carbon-containing fuels. It is also consumed when oxidizable minerals undergo chemical weathering, such a

$$4\,FeO + O_2 \rightarrow 2\,Fe_2O_3$$

All of the oxygen in the atmosphere was originally placed there by photosynthesis shown by

$$CO_2 + H_2O + h\nu \rightarrow (CH_2O) + O_2$$

where $\{CH_2O\}$ is a generic formula representing biomass

Unlike molecular oxygen, which can undergo direct photodissociation in the stratosphere, the very stable N_2 molecule does not encounter ultraviolet radiation sufficiently energetic to cause its photodissociation at altitudes below 100 km. However, nitrogen dioxide, NO_2, readily undergoes photodissociation in the troposphere.

$$NO_2 + h\nu \rightarrow NO + O$$

to generate highly reactive O atoms. These in turn can attack hydrocarbons in the atmosphere, leading to the formation of photochemical smog discussed later in this chapter.

This page titled 2.5.10.1: More Than Just Air to Breathe is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.1: More Than Just Air to Breathe** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.2: The Gas Laws

The behavior of gases in the atmosphere is governed by several fundamental **gaslaws** which are covered briefly here. In using these laws, it should be kept in mind that the quantity of gas is most usefully expressed in numbers of moles. There are many units of pressure, but the most meaningful conceptually is the **atmosphere** (atm) where 1 atmosphere is the average pressure of air in the atmosphere at sea level. (Air has pressure because of the mass of all the molecules of air pressing down from the atmosphere above; as altitude increases, this pressure becomes less.) For calculations involving temperature, the **absolute** temperature scale is used in which each degree is the same size as a degree Celsius (or Centigrade, the temperature scale used for scientific measurements and for temperature readings in most of the world), but zero is 273 degrees below the freezing point of water, which is taken as zero on the Celsius scale. Three important gas laws are the following:

> **Avogadro's law:** At constant temperature and pressure the volume of a gas is directly proportional to the number of moles; doubling the number of moles at a constant temperature and pressure doubles the volume.
>
> **Charles' law:** At constant pressure the volume of a fixed number of moles of gas is directly proportional to the absolute temperature (degrees Celsius +273) of the gas; doubling the absolute temperature at constant pressure doubles the volume.**Boyle's law:**
>
> At constant temperature the volume of a fixed number of moles of gas is inversely proportional to the pressure; doubling the pressure halves the volume.

These three laws are summarized in the **general gas law** relating volume (V), pressure (P), number of moles (n), and absolute temperature (T) expressed as

$$\mathrm{PV} = n\,\mathrm{RT}$$

where R is a constant.

Mathematical calculations involving the gas laws are simple. One of the most common such calculations is that of changes in volume resulting from changes in pressure, temperature, or moles of gas. The parameter that does not change is the constant R. Using subscripts to represent conditions before and after a change yields the following relationship:

$$R = \frac{P_1 V_1}{n_1 T_1} = \frac{P_2 V_2}{n_2 T_2}$$

This equation can be arranged in a form that can be solved for a new volume resulting from changes in P, n, or T

$$V_2 = V_1 \times \frac{n_2 T_2 P_1}{n_1 T_1 P_2}$$

As an example, calculate the volume of a fixed number of moles of gas initially occupying 12.0 liters when the temperature is changed from 10°C to 90°C at constant pressure. In order to use these temperatures, they must be changed to absolute temperature by adding 273°. Therefore, T_1=10° + 273° = 283° , and T_2= 90° + 273° = 363°. Since n and P remain constant, they cancel out of the equation yielding

$$V_2 = V_1 \times \frac{T_2}{T_1} = 12.0L \times \frac{363^\circ}{283^\circ} = 15.4L$$

As another example consider the effects of a change of pressure, holding both the temperature and number of moles constant. Calculate the new volume of a quantity of gas occupying initially 16.0 L at a pressure of 0.900 atm when the pressure is changed to 1.20 atm. In this case, both n and T remain the same and cancel out of the equation giving the following relationship:

$$V_2 = V_1 \times \frac{P_1}{P_2} = 16.0L \times \frac{0.900\ \mathrm{atm}}{1.20\ \mathrm{atm}} = 12.0L$$

Note that an increase in temperature *increases* the volume and an increase in pressure *decreases* the volume.

This page titled 2.5.10.2: The Gas Laws is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.2: The Gas Laws** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.3: The Protective Atmosphere

The atmosphere is the air around and above us. We know we must have air to breathe. A human deprived of air's life-giving oxygen for just a brief time will lose consciousness, and within a few minutes will die. But air is far more than just a source of oxygen. That is because it protects Earth's organisms in ways that are absolutely essential for their existence. One major protective function is to act as a blanket to keep us warm. It does that by reabsorbing the infrared radiation by which Earth radiates the energy that it receives from the sun. By delaying the exit of this energy into outer space, the average temperature of Earth's surface remains at about 15°C at sea level, though much colder at certain times and places and significantly warmer at others. Without this warming effect, plants could not grow and most other known organisms could not exist. The second protective function of the atmosphere is absorption of very short wavelength ultraviolet solar radiation. Were this radiation to reach our level, it would tear apart biomolecules, making it impossible for most life forms to exist.

Although one might get the impression that the atmosphere is very thick, it is "tissue thin" compared to Earth's diameter. Consider a corporate jet aircraft cruising at 35,000 feet (about 6.6miles or 10.7 kilometers). In the unlikely event of sudden, catastrophic loss of pressure in the pressurized cockpit, the pilot has only about *15seconds* to grab an oxygen mask before losing consciousness (the passengers in the cabin have an equally short time, but it is more important for the pilot to stay conscious and dive to a lower altitude). The reason for this is that virtually all the air in the atmosphere is below the altitude of around 11 km. By way of comparison, Earth'sdiameter is almost 13,000 km.

The altitude at which high-flying jet aircraft cruise marks the upper limit of the lowest of several layers of the atmosphere, the **troposphere**, which extends from sea level to about 11 km (Figure 10.1). As anyone who has driven to the top of Pike's Peak or some other mountain knows, the troposphere gets cooler with increasing altitude, from an average temperature of 15°C at sea level to an average at 11 km of -56°C. Above the layer of the troposphere, however, atmospheric temperature increases to an average of -2°C at 50 km altitude. The layer above the troposphere is the **stratosphere**, which is heated by the absorption of intense ultraviolet radiation from the sun(Figure 10.2). There is virtually no water vapor in the stratosphere, and it contains ozone, O3, and O atoms as the result of ultraviolet radiation acting upon stratospheric O2. Beyond the stratosphere are layers called the mesosphere and thermosphere, but they are relatively less important in the discussion of the atmosphere.

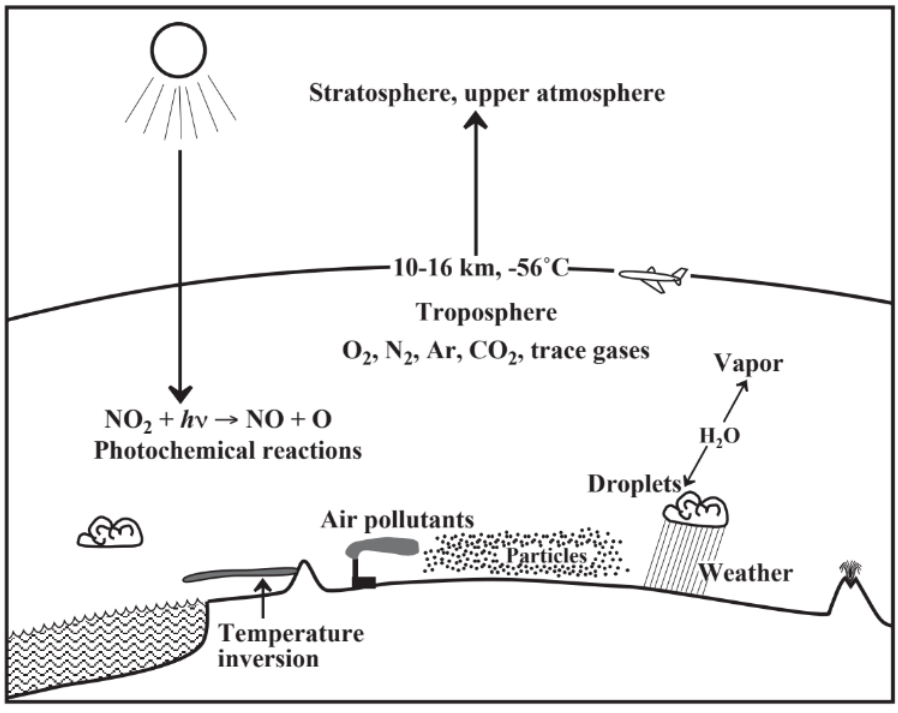


Figure 10.1. The troposphere is the very thin layer of the atmosphere closest to Earth, containing most of the atmosphere's air and water vapor. It is the source of oxygen, carbon dioxide, nitrogen, and water used by living organisms and as raw materials for manufacturing. With the important exception of stratospheric ozone destruction, it is where most air pollution phenomena occur

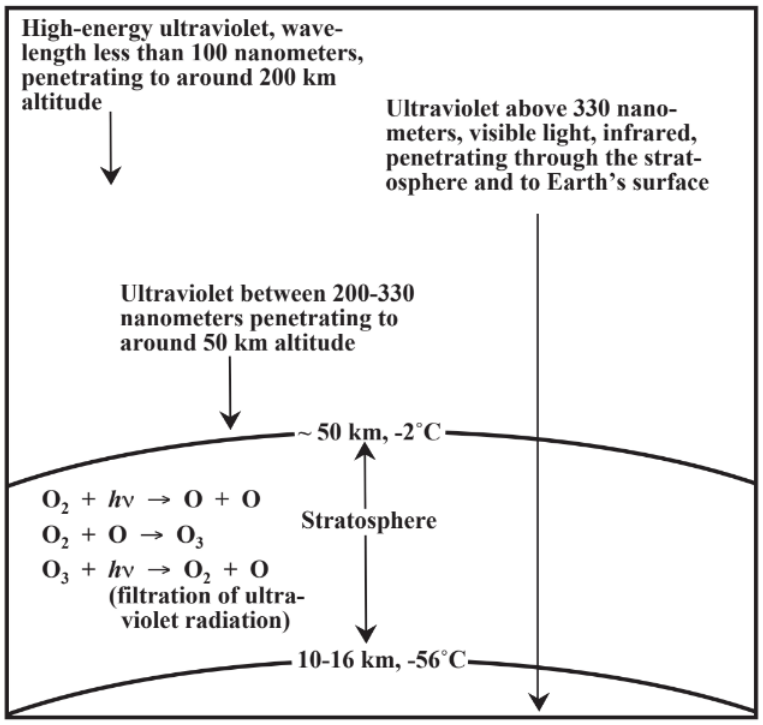


Figure 10.2. The upper atmosphere including the stratosphere and regions beyond is particularly important in the absorption of radiation that would make life impossible if it reached Earth's surface. The layer of ozone, O3, in the stratosphere is of the utmost importance and one that is subject to damage from anthropogenic species released into the atmosphere

---

This page titled 2.5.10.3: The Protective Atmosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.3: The Protective Atmosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.4: Atmospheric Chemistry and Photochemical Reactions

Atmospheric chemistry describes chemical processes that occur in the atmosphere. One notable aspect of atmospheric chemistry is that it occurs in the gas phase where molecules are relatively far apart, so a molecule or a fragment of a molecule may travel some distance before bumping into another species with which it reacts. This is especially true in the highly rarefied regions of the stratosphere and above.

A second major aspect of atmospheric chemistry is the occurrence of **photochemicalreactions** that are initiated when a photon (essentially a packet of energy associated with electromagnetic radiation) of ultraviolet radiation is absorbed by a molecule. The energy of a photon, E, is given by E =*h*ν where *h* is Planck's constant andvis the frequency of the radiation. Electromagnetic radiation of a sufficiently short wavelength breaks chemical bonds in molecules, leading to the formation of reactive species that can participate in reaction sequences called **chain reactions.**

An example of an important chain reaction sequence that begins with photochemical dissociation of a molecule is the one that occurs when chlorofluorocarbons get into the stratosphere. Chlorofluorocarbons are given the trade name of Freons and consist of carbon atoms to which are bonded fluorine and chlorine atoms. Noted for their extreme chemical stability and low toxicities, they were once widely used as refrigerant fluids in air conditioners, as aerosol propellants for products such as hair spray, and for foam blowing to make very porous plastic or rubber foams. Dichlorodifluoromethane, CCl_2F_2, was used in automobile air conditioners. Released to the atmosphere, this compound remained as a stable atmospheric gas until it got to very high altitudes in the stratosphere. In this region, ultraviolet radiation of sufficient energy (*h*ν)is available to break the very strong C-Cl bonds,

$$CCl_2F_2 + h\nu \rightarrow \cdot CClF_2 + Cl \cdot$$

releasing Cl atoms. The dot represents a single unpaired electron remaining with the Cl atom when the bond in the molecule breaks. Species with such unpaired electrons are very reactive and are called **free radicals**. As discussed in Chapter 2 Section 2.13 and shown by reactions 2.13.1 and 2.13.2, in addition to molecular O_2 there are oxygen atoms and molecules of ozone, O_3, also formed by photochemical processes in the stratosphere. A chlorine atom produced by the photochemical dissociation of CCl_2F_2 as shown in Reaction 10.4.1 can react with a molecule of O_3 to produce O_2 and another reactive free radical species, ClO•. This species can react with free O atoms which are present along with the ozone to regenerate Cl atoms, which in turn can react with more O_3 molecules. These reactions are shown below:

$$Cl \cdot + O_3 \rightarrow O_2 + ClO \cdot$$

$$ClO \cdot + O \rightarrow O_2 + Cl \cdot$$

These are chain reactions in which ClO• and Cl• are continually reacting and being regenerated, the net result of which is the conversion of O_3 and O in the atmosphere to O2. One Cl atom can bring about the destruction of as many as 10,000 ozone molecules! Ozone serves a vital protective function in the atmosphere as a filter for damaging ultraviolet radiation, so its destruction is a very serious problem that has resulted in the banning of chlorofluorocarbon manufacture.

Very small particles of the size of a micrometer or less called **aerosols** are important in atmospheric chemical processes. Photochemical reactions often result in the production of particles. Particle surfaces can act to catalyze (bring about) atmospheric chemical reactions. Some particles in the atmosphere consist of water droplets with various solutes dissolved in them. Solution chemical reactions can occur in these droplets. One such process is believed to be the conversion of gaseous atmospheric sulfur dioxide (SO_2) to droplets of dilute sulfuric acid (H_2SO_4), which contribute to acid rain. Some very small particles, such as sea salt crystals entrained into the atmosphere by wind-blown seawater spray droplets and formed by evaporation of water from the droplets, act as **condensationnuclei** around which raindrops form.

The Ionosphere

An important kind of photochemical reaction that occurs at altitudes generally above the stratosphere (50 km and higher) is the formation of ions by the action of ultraviolet and cosmic radiation energetic enough to remove electrons (e^-) from molecules as shown by the example below:

$$N_2 + h\nu \rightarrow N_2^+ + e^-$$

The ions formed are very reactive, but air is so rarefied at the altitudes at which they form that they persist for some time before reverting to neutral species. This results in an atmospheric layer called the **ionosphere** in which ions are constantly being formed and neutralized. At night when the solar radiation responsible for ion formation is shielded by Earth, the predominant process is recombination of positive ions with electrons, a phenomenon that proceeds most rapidly in the lower, denser regions of the ionosphere. The result is a lifting of the ionosphere, a phenomenon first hypothesized in 1901 when Marconi, attempting to bridge the Atlantic ocean with shortwave radio discovered that radio waves could be propagated over long distances, especially at night, making long distance shortwave radio transmission possible. For a time in the 1900s until made obsolete by satellite and fiber optics, the ionosphere was a useful part of the atmosphere's natural capital (see Section 10.12) by making possible long-distance shortwave radio broadcasts.

This page titled 2.5.10.4: Atmospheric Chemistry and Photochemical Reactions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.4: Atmospheric Chemistry and Photochemical Reactions** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.5: The Enormous Importance of Climate

The most important aspect of the atmosphere for humans and other living beings is **climate**. Climate consists of long-term trends in weather and varies a lot over Earth's surface. For example, the climate in desert regions of the world may be relatively hot and dry, but the weather in such regions may at times produce torrential rainfall or frigid temperatures. Of all the things that humans might do to irreversibly and catastrophically harm Mother Earth, the only home that they have or ever can have, the potential to change climate is the most serious. It is well established that climate has changed markedly in past times, most notably during several ice ages that each lasted for thousands of years. Volcanic eruptions noted in historical accounts have caused temporary cooling of the atmosphere and widespread hunger when summer growing seasons were cut short as a result. Tree-ring data indicate centuries-long droughts in parts of the world in the past, which resulted in the decline of some civilizations in the present-day U.S. Southwest, for example. Although these were natural events, recent weather data indicate changes in microclimate due to human activities. For example, some urban areas in Southeast Asia have become darker in recent decades due to particulate air pollution.

Of all the things that humans may be doing that could change climate, emissions of large quantities of gases that cause atmospheric warming are the most serious. Of these, carbon dioxide, CO_2, is the most important. Carbon dioxide is a normal constituent of the atmosphere, essential as a source of carbon for plant photosynthesis. Along with water vapor and other trace gases, atmospheric carbon dioxide absorbs outgoing infrared radiation from Earth, thus keeping the planet's surface temperature at a tolerable level. Levels of carbon dioxide gas in the atmosphere are now about 390 parts per million by volume. This represents an almost 50% increase overestimated pre-industrial concentrations of 260 ppm. Furthermore, as shown by the plot in Figure 10.4, global CO_2 levels are increasing by about 1 ppm per year, and ice core evidence indicates that these levels were only about 200 ppm at the peak of the last ice age around 18,000 years ago. So, humans are clearly increasing atmospheric carbon dioxide levels significantly, largely through the combustion of carbon-containing fossil fuels and as the result of destruction of forests. The importance of photosynthesis in determining atmospheric carbon dioxide is illustrated by the inset in Figure 10.4 showing an annual fluctuation of about 5 ppm CO_2 in the northern hemisphere attributed to photosynthesis. The minimum in this cycle occurs around September at the end of the summer growing season and the maximum occurs around April as photosynthesis is getting underway after winter.

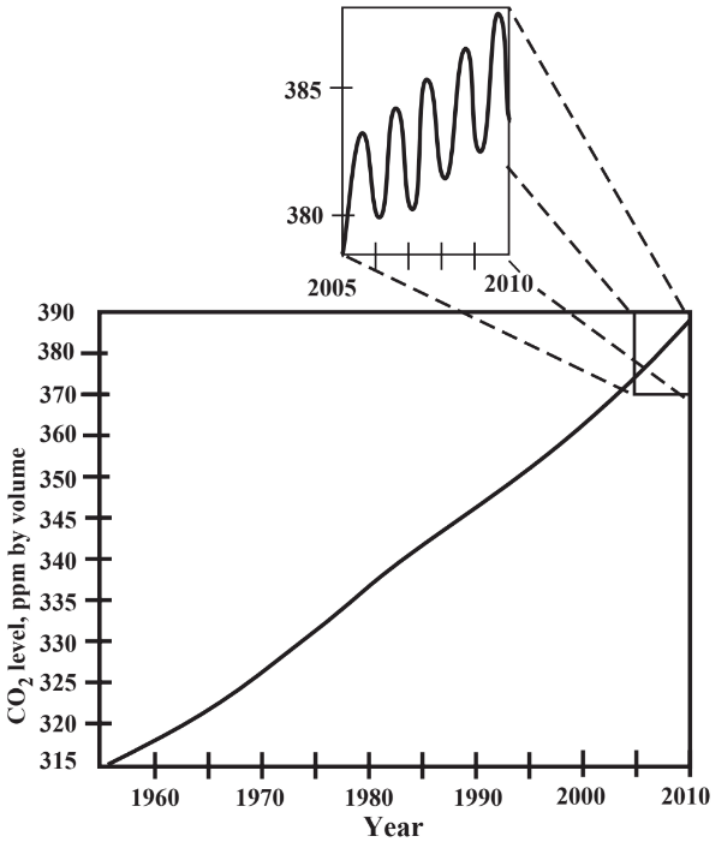


Figure 10.4. Concentrations of atmospheric CO_2 levels in recent decades. These levels are increasing by approximately 1 part per million by volume each year

The concern with increasing carbon dioxide in the atmosphere is that it will lead to — indeed *is* leading to — an excess of a good thing, warming of the global atmosphere. This is the now well known **greenhouse effect**, which may not become as dire as some experts predict, but which has a real possibility of becoming the worst environmental problem so far created by humans.

Is Earth warming? Sophisticated computer models indicate that it is backed by evidence from increasingly accurate temperature records over more than 100 years. These temperature records have been especially accurate over the last several decades because they have been read over all Earth's surface by satellite. Global temperatures analyzed by the Goddard Institute for Space Studies (GISS) in New York City show that the 1980s were the warmest decade on record globally since reasonably accurate global temperatures have been measured. The 1980s were followed by a warmer decade in the 1990s and the warmest decade of all from the beginning of 2000 to the end of 2009 (see Figure 10.5). The hottest year ever documented historically was 2005 with 1998, 2002, 2003, 2006, 2007, and 2009 all essentially tied for second.

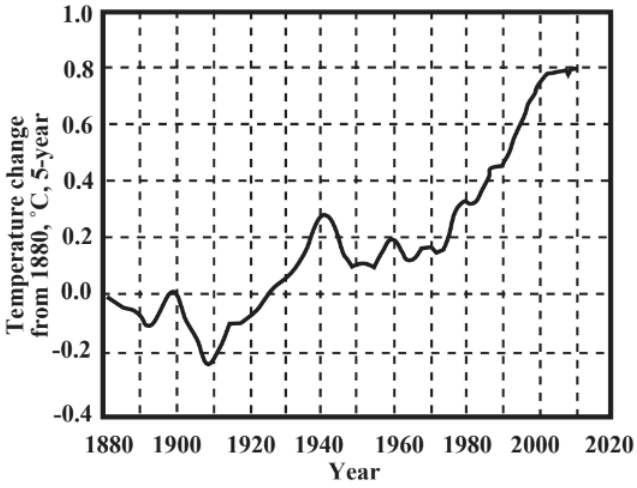


Figure 10.5. Global temperatures since 1880. Since about 1970 these measurements have been particularly accurate because they were taken by scanning throughout the globe by satellite

Gases other than carbon dioxide may be involved in greenhouse warming. These include chlorofluorocarbons and N_2O. The one most likely to cause a problem is methane, CH4, which has increased from estimated pre-industrial atmospheric levels of 0.70 ppm, to present values of l.8ppm. Although these values are much lower than those of carbon dioxide, each methane molecule is 20–30 times more effective in trapping heat than is each CO_2 molecule. A number of humana ctivities have contributed to the release of methane. Part of this is due to leakage of natural gas, which consists of methane, and from release as a byproduct of petroleum production. The 2010 blowout of the Deepwater Horizon well in the Gulf of Mexico released large quantities of methane that had been sequestered largely in combination with water as solid methane hydrates far below the ocean floor. These methane-containing structures belong to a class of materials called clathrates, an inherently unstable network of host molecules containing open cavities. In the case of methane hydrates the clathrates are clusters of approximately 46 H_2O molecules that enclose molecules of CH_4. Methane hydrates can form and be stable under pressure at temperatures significantly above the freezing point of water. Estimates are that more methane is contained in these structures globally than in all the deposits of natural gas. A major concern with global warming is that methane contained in methane hydrates held by permafrost will be released causing a feedback loop that will put even more methane into the atmosphere and result in even greater warming. Bacteria growing in the absence of oxygen in municipal refuse landfills, in rice paddies, and in the stomachs of ruminant animals (cows, sheep, moose) release enormous quantities of methane.

Although there could be some benefits of mild global warming, the net effect would almost certainly be bad, perhaps catastrophic. Climate models predict an average global temperature increase of 1.5–5°C. That does not sound like much, but it is about as much again as the temperature increase that occurred from the last ice age until now. Especially if the warming is toward the high side of the projected range, it would greatly affect climate and rainfall. The melting of the polar and Greenland ice caps along with expansion of warmer ocean water would cause sea levels to rise as much as 0.5–1.5 meters. Decreased rainfall and increased water evaporation would contribute to severe drought and water shortages that could make some currently popular areas of the world virtually uninhabitable.

Can Green Chemistry Help Deal With Global Warming?

Green chemistry and the related area of industrial ecology can help deal with the problem of global warming in two major respects. The first approach is to provide means to prevent global warming from taking place. The second approach is in coping with global warming, if it occurs.

The prevention of global warming is best accomplished by avoiding the release of potential greenhouse gases. As noted above, the most significant of these is carbon dioxide. One way to reduce the release of carbon dioxide is by using biomass as fuel or raw material for the manufacture of various products. Burning a biomass fuel does release carbon dioxide to the atmosphere, but an exactly equal amount of carbon dioxide was removed from the atmosphere in the photosynthetic process by which the biomass was made, so there is no net addition of CO_2. Unless or until biomass-derived materials used in feedstocks are burned or biodegraded, their use represents a net loss of carbon dioxide from the atmosphere.

Another potential use of green chemistry to prevent addition of carbon dioxide to the atmosphere is through **carbonsequestration** in which carbon dioxide is produced, but is bound in a form such that it is not released to the atmosphere. This approach has the greatest potential in applications where the carbon dioxide is produced in a concentrated form. In Chapter 16, reactions are shown by which carbon from coal is reacted with oxygen and water to produce elemental hydrogen and carbon dioxide. The net reaction for this production is the following:

$$2\,C + O_2 + 2\,H_2O \rightarrow 2\,CO_2 + 2\,H_2$$

The hydrogen generated can be used as a pollution-free fuel in fuel cells or combustion engines. The carbon dioxide can be pumped into deep ocean waters, although this has the potential to lower ocean pH slightly, which would be detrimental to marine organisms. Another option is to pump the carbon dioxide deep underground. A side benefit of the latter approach is that in some areas carbon dioxide pumped underground can be used to recover additional crude oil from depleted oil-bearing formations.

An indirect green chemistry approach to the reduction of carbon dioxide emissions is to develop alternative methods of energy production. One thing that would be very beneficial is the development of more efficient photovoltaic cells. These devices have become marginally competitive for the generation of electricity, and even relatively small improvements in efficiency would enable their much wider use, replacing fossil fuel sources of electricity generation. Another device that would be extremely useful is a system for the direct photochemical dissociation of water to produce elemental hydrogen and oxygen, which could be used in fuel cells. An application of green biochemistry that would reduce carbon dioxide emissions is the development of plants with much higher efficiencies for photosynthesis. Plants now are only about 0.5%efficient in converting light energy to chemical energy. Raising this value to only 1% would make a vast difference in the economics of producing biomass as a substitute for fossil carbon.

Green chemistry can also be applied in the prevention of release of greenhouse gases other than carbon dioxide, especially ultrastable volatile compounds that have a high greenhouse gas potential. An excellent example of green chemistry has been the replacement of chlorofluorocarbons (Freons) with analogous compounds having at least one C-H bond, that are rather readily destroyed in the troposphere. Although this was done to prevent destruction of stratospheric ozone by chlorofluorocarbons, it has been useful to reduce greenhouse warming. Both kinds of compounds act as greenhouse gases, but those with at least one C-H bond last for much shorter times during which they are available to absorb infrared radiation. As discussed inSection 10.10, green chemistry can be applied to avoid generating extremely stable sulfurhexafluoride, SF_6, and completely fluorinated hydrocarbons, such as CF_4.

Another approach is to limit the emissions of methane, CH_4. Large quantities of methane are released by anoxic bacteria growing in flooded rice paddies. By developing strains of rice and means of cultivation that enable the crop to be grown on unflooded soil, this source of methane can be greatly reduced. Methane collection systems placed in municipal waste landfills can prevent the release of methane from this source and provide a source of methane fuel.

Green chemistry, biochemistry, and biology can be used to deal with global warming when it occurs. Crops, fertilizers, and pesticides can be developed that enable plants to grow under the drought conditions that would follow global warming. Another approach is the development of salt-tolerant crops that can be grown on soil irrigated with saline water, where fresh water supplies are limited.

This page titled 2.5.10.5: The Enormous Importance of Climate is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.6: The Enormous Importance of Climate** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.6: Atmospheric Particle Pollutants

Although gases that cause global warming are arguably the most serious air pollutants, other air pollutants can cause serious problems with the atmospheric environment. These are discussed in the remainder of this chapter.

Because of their ability to reduce visibility and light, atmospheric particles are the most visible form of air pollution. Commonly called **particulates** in air pollution parlance, atmospheric *aerosols* are solids or liquids less than 100 micrometers (millionths of a meter,μm) in diameter, and commonly in a size range of 0.001 to 20μm. They may be inorganic or organic materials and may belong to the two general classes of **dispersionaerosols** formed by grinding solids, dispersing dusts, or atomizing liquids, and **condensationaerosols** produced by the condensation of gases or vapors, often formed as the result of atmospheric chemical processes. Common dispersion aerosols include water droplets from sea spray, solid particles of NaCl left over when water evaporates from sea spray droplets, cement dust, soil dust dispersed by wind, foundry dust, and pulverized coal. Carbon black, metal fumes, and combustion nuclei form as condensation aerosols from combustion or partial combustion. Liquid particle **mists** include raindrops, fog, cloud droplets, and droplets of sulfuric acid produced when atmospheric SO_2 is oxidized. Organisms produce an abundance of particles. For those afflicted with allergies, the most annoying such particles are plant pollen. Other particles of biological origin include viruses, bacteria, and spores of bacteria and fungi.

In the past and even now in some areas of the world, one of the more troublesome sources of atmospheric particles was **fly ash**, a byproduct residue from combustion of liquids or very finely divided coal. Often the most abundant component of fly ash is elemental carbon left over from incompletely burned fuel. Fly ash commonly includes oxides of aluminum, calcium, iron, and silicon, as well as some magnesium, sulfur, titanium, phosphorus, potassium, and sodium. With properly operating emission control devices, fly ash emissions are now well controlled.

One health concern with particles, especially those from combustion sources, is their ability to carry toxic metals. Of these, lead is of the greatest concern because it usually comes closest to being at a toxic level. Problems with particulate lead in the atmosphere have been greatly reduced by the elimination of tetraethyllead as a gasoline additive, an application that used to spew tons of lead into the atmosphere every day. Another heavy metal that causes considerable concern is mercury, which can enter the atmosphere bound to particles or as vapor-phase atomic mercury. Airborne mercury from coal combustion can become a serious water pollution problem leading to unhealthy accumulations of this toxic element in some fish. Other metals that can cause problems in particulate matter are beryllium, cadmium, chromium, vanadium, nickel, and arsenic (a metalloid).

Some areas of the world, including parts of the United States, have problems with radioactive particles resulting from underground sources of radioactive **radon**, a noble gas product of radium decay. The two major radon isotopes, 222Rn (half-life 3.8 days) and 220Rn (half-life 54.5 seconds)are alpha particle emitters that decay to radioactive 218Po and 216Po, respectively. These radionuclides are nongaseous and adhere readily to atmospheric particulate matter, which, along with gaseous radon, can cause significant indoor air pollution and potential health problems.

Pollutant particles in the atmosphere have both direct and indirect effects. The most obvious direct effects are reduction and distortion of visibility. The light scattering effects of particles in a size range of 0.1μm–1μm are especially pronounced due to interference phenomena resulting from the particles being about the same size as the wavelengths of visible light. Particles also have direct health effects when inhaled. This is especially true of very small particles that can be carried into the innermost parts (alveoli) of lungs. An indirect effect of particles is their ability to serve as reaction sites for atmospheric chemical reactions. They also act as nucleation bodies upon which water vapor condenses.

Limiting Particulate Emissions

The first widespread measures to limit air pollution were directed at control of particle emissions. These measures have become very effective so that the “smoke” that one sees emanating from smokestacks usually consists of droplets of water formed by condensation of steam.

The simplest method of particle control from stack gas and other gases released to the atmosphere consists of **sedimentation** in which particles entrained in stack gas are allowed to settle by gravity in relatively large chambers. Sedimentation is most effective for larger particles. **Inertial mechanisms** operate by spinning a gas in a round chamber such that particles impinge upon the container walls by centrifugal force. **Fabric filters** contained in **baghouses** act to filter particles from air or stack gas (Figure 10.6). The mechanism employed provides for periodic shaking of the fabric filters to collect particles held on their walls, thus restoring

gas flow through the fabric. Numerous factors including moisture levels, particle abrasion, particle size, and acidity or alkalinity of the gas and particles must be considered in choosing filter fabrics. **Scrubbers** that spray water or solutions into stack gas are employed to literally wash particles out of gas. In some cases these are operated with a minimal amount of water, which evaporates, so that a solid material is collected. One of the most effective means of particle control consists of **electrostatic precipitators**. These devices use a very high voltage to impart a negative charge onto particles from a central electrode, and the particles are attracted to, and collect on the positively charged walls of the precipitator.

In keeping with the practice of sustainability the particulate matter, such as that collected by a fabric filter in a baghouse, may be used for various purposes. Typically, particulate matter from lead or zinc smelting operations is recycled back into the metal recovery process. Lime kiln dust is often used as agricultural lime. Some kinds of coal fly ash could be used as a source of aluminum if aluminum ore (bauxite) becomes scarce.

This page titled 2.5.10.6: Atmospheric Particle Pollutants is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.7: Atmospheric Particle Pollutants** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.7: Pollutant Gaseous Oxides

Carbon Monoxide

Carbon dioxide, CO_2, which is an essential ingredient of the atmosphere, but a potential cause of global warming at elevated levels, was discussed in some detail in Section 10.6. Carbon monoxide, CO, is an air pollutant of some concern because of its direct toxicity to humans. Carbon monoxide is toxic because it binds to blood hemoglobin and prevents the hemoglobin from transporting oxygen from the lungs to other tissues. Global and regional levels of atmospheric carbon monoxide are too low to be of concern. However, local levels in areas with heavy automobile traffic can become high enough to pose a health hazard and on some congested urban streets have reached levels of 50–100 ppm. The use of exhaust pollution control devices on automobiles have lowered these levels significantly during the last 30 years. The numerous fatal cases of carbon monoxide poisoning that occur each year are almost always the result of improperly vented heating devices in indoor areas.

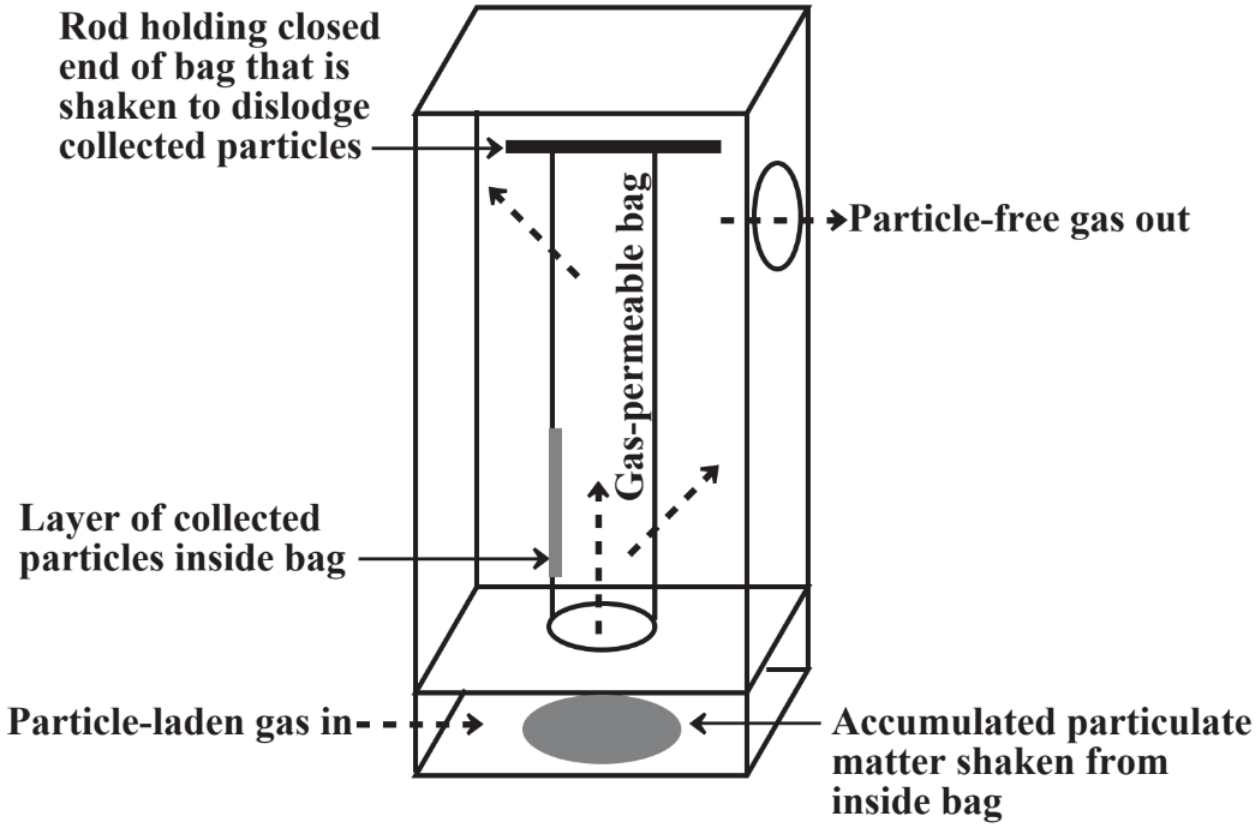


Figure 10.6. Illustration of a unit of a single bag in a baghouse for removal of particulate matter from a gas stream. The flow of gas is illustrated by the dashed lines. A bag is mounted on an opening in a plate such that particle-laden gas is forced up through the inside of the cylindrical bag and is filtered through the bag walls, then flows out as filtered gas. The particulate matter accumulates on the inside of the bag. Periodically, the gas flow through the bag is stopped and the rod suspending the closed upper end of the bag is shaken to dislodge particulate matter from the bag walls, which is collected at the bottom of the apparatus.

Carbon monoxide is produced by partial combustion of fuels, largely in the internal combustion engine. Carbon monoxide emissions can be reduced by careful control of combustion conditions. Running the exhaust along with pumped air over an exhaust catalyst can oxidize carbon monoxide to nontoxic carbon dioxide

$$2\,CO + O_2 \rightarrow 2\,CO_2$$

Modern automobile engines use precise computerized control of engine operating parameters along with exhaust catalysts to control carbon monoxide emissions.

Sulfur Dioxide

Sulfur dioxide enters the atmosphere as the result of the following:

- Direct emissions from volcanoes
- Atmospheric oxidation of H_2S emitted to the atmosphere by bacteria and from geothermal sources (volcanoes, hot springs, geysers)
- Atmospheric oxidation of dimethyl sulfide, $(CH_3)_2S$, emitted to the atmosphere from marine organisms
- Pollutant sources from the combustion of organic sulfur and iron pyrite, FeS_2, in fossil fuels

The pollutant sources are of most concern because of their contribution to local and regional air pollution problems and because they are sources that humans can do something about.

The fate of sulfur dioxide in the atmosphere is oxidation and reaction with water to produce sulfuric acid. The overall process is complex and not completely understood, but it can be described by the reaction.

$$2\,SO_2 + O_2 + 2\,H_2O \rightarrow 2\,H_2SO_4$$

This process is generally rather slow in the atmosphere, but it can be quite rapid under conditions of photochemical smog formation (see Section 10.11) in which highly reactive oxidizing species are present. It is very important because it is the main mechanism for forming acid rain, which can be directly harmful to vegetation, fish (especially fingerlings), and materials, such as building stone that can be attacked by acid. Sulfur dioxide forms aerosol droplets of sulfuric acid in the atmosphere. As a result, much of the Eastern United States is covered by a slight haze of sulfuric acid droplets during much of the year. In recent years, some volcanic eruptions have blasted enough sulfur dioxide into the atmosphere to produce a sufficient amount of sunlight-reflecting sulfuric acid aerosol to cause a noticeable cooling of the atmosphere.

In addition to indirect effects from the formation of acid rain, sulfur dioxide affects some plants directly, causing leaf necrosis (death of leaf tissue). Another symptom of sulfur dioxide phytotoxicity (toxicity to plants) is chlorosis, a bleaching or yellowing of green leaves.

The most straightforward means of reducing sulfur dioxide emissions is to avoid having sulfur in fuels. As discussed below, sulfur compounds are removed from natural gas and petroleum. Coal often has high levels of sulfur, and during recent years there has been a major shift to low sulfur coal in power plants. Much of the pyritic sulfur, FeS_2, can be washed from coal because it exists in a separate mineral phase that can be separated from the combustible organic matter in coal. However, about half of the sulfur in coal typically is bound to the coal as organic sulfur, and cannot be removed.

A number of coal-fired power plants have installed systems for removing sulfur dioxide resulting from the combustion of coal. One such approach uses **fluidized bed combustion** in which pulverized coal is blasted into a hot bed of calcium oxide, where the coal is burned, and sulfur dioxide is bound by the following reaction.

$$CaO + SO_2 \rightarrow CaSO_3$$

Some of the $CaSO_3$ product is oxidized to $CaSO_4$. Another approach uses a slurry of calcium hydroxide (lime, $Ca(OH)_2$) to react with sulfur dioxide in stack gas.

$$Ca(OH)_2 + SO_2 \rightarrow CaSO_3 + H_2O$$

Although effective in removing sulfur dioxide, this process uses enormous quantities of limestone ($CaCO_3$) as a source of lime and produces huge quantities of byproduct.

Green Chemistry and Sulfur Dioxide

The problem of sulfur in fuel provides an excellent illustration of the potential for the application of green chemistry to the elimination of a pollution problem. Sulfur is a valuable raw material required in the manufacture of sulfuric acid, one of the largest volume chemicals made. Hydrogen sulfide, H_2S, occurs in large quantities in natural gas, such as that produced in the Canadian province of Alberta. This hydrogen sulfide must be removed from the natural gas. Rather than presenting a pollution problem, it is converted to elemental sulfur, then used to make sulfuric acid.

Another green chemistry approach to the reclamation of waste sulfur is practiced in Kalundborg, Denmark, the site of the world's most clearly recognizable system of industrial ecology (see Chapter 14). The huge coal-fired power plant in Kalundborg uses lime scrubbing to remove sulfur dioxide from stack gas. The calcium sulfite product of this process is oxidized,

$$CaSO_3 + \tfrac{1}{2}\,O_2 + 2\,H_2O \rightarrow CaSo_4 \cdot 2\,H_2O$$

to generate gypsum, $CaSO_4$•$2H_2O$. This mineral is then used to make wallboard, thus solving a pollution problem from the production of spent lime and a raw materials problem arising from the need for gypsum to make wallboard needed for building construction.

Nitrogen Oxides

Nitrous oxide (N_2O), colorless, odorless, nitric oxide (NO), and pungent-smelling, red-brown nitrogen dioxide (NO_2) occur in the atmosphere. Of these, nitrous oxide is generated by bacteria and its release is one of the ways in which chemically fixed nitrogen in the biosphere is returned to the atmosphere. It is not involved much with chemical processes in the troposphere, but undergoes photochemical dissociation in the stratosphere:

$$N_2O + h\nu \rightarrow N_2 + O$$

Both NO and NO_2, collectively designated as NO_x, are produced from natural sources, such as lightning and biological processes, and from pollutant sources. Pollutant concentrations of these gases can become too high locally and regionally, causing air pollution problems. A major pollutant source of these gases is the internal combustion engine in which conditions are such that molecular elemental nitrogen and oxygen react,

$$N_2 + O_2 \rightarrow 2\,NO$$

to produce NO. Combustion of fuels that contain organically bound nitrogen, such as coal, also produces NO. Atmospheric chemical reactions convert some of the NO emitted to NO_2.

Exposed to electromagnetic radiation of wavelengths below 398 nm, nitrogen dioxide undergoes photodissociation,

$$NO_2 + h\nu \rightarrow NO + O$$

to produce highly reactive O atoms. The O atoms can participate in a series of chain reactions through which NO is converted back to NO_2, which can undergo photodissociation again to start the cycle over. Nitrogen dioxide is very reactive, undergoing photodissociation within a minute or two in direct sunlight.

Nitrogen dioxide, NO_2, is significantly more toxic than NO, although concentrations of NO_2 in the outdoor atmosphere rarely reach toxic levels. Accidental releases of NO_2 can be sufficient to cause toxic effects or even death. Brief exposures to 50–100 parts per million (ppm) of NO_2in air inflames lung tissue for 6-8 weeks followed by recovery. Exposure to 500 ppm or more of NO_2 causes death within 2–10 days. Exposure to 100-500 ppm of NO_2 causes a lung condition with the ominous name of *bronchiolitisfibrosa obliterans* that is fatal within 3–5 weeks after exposure. Fatal incidents of NO_2poisoning have resulted from accidental release of the gas used as an oxidant in rocket fuels and from burning of nitrogen-containing celluloid and nitrocellulose moving picture film (nitrocellulose has been used as an explosive and has long been banned in film because of catastrophic fires that killed numerous people). Plants exposed to nitrogen dioxide may suffer decreased photosynthesis, leaf spotting, and breakdown of plant tissue.

The release of NO from combustion sources can be reduced by limiting excess air so that there is not enough excess oxygen to produce NO according to Reaction 10.8.7. The production of NO is also reduced by relatively lower combustion temperatures. Exhaust catalytic converter reduce NO_x emissions from automobile exhausts. Scrubbing NO_x from furnace and power plant stack gases is difficult due to the low water solubilities of NO_x gases.

This page titled 2.5.10.7: Pollutant Gaseous Oxides is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.8: Pollutant Gaseous Oxides** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.8: Acid Rain

Along with hydrogen chloride, HCl, emitted to the atmosphere by the combustion of chlorine-containing organic compounds, sulfur dioxide and nitrogen oxides react in the atmosphere to produce strongly acidic H_2SO_4 and HNO_3, respectively. Incorporated into rainwater, these acids fall to the ground as **acid rain**. A more general term, **acid deposition**, refers to the effects of atmospheric strong acids, acidic gases (SO_2), and acidic salts ($NH_4NO_3$and NH_4HSO_4). Acid deposition is a major air pollution problem.

Figure 10.7 shows a typical distribution of acidic precipitation in the 48 contiguous U.S. states. This figure illustrates that acidic precipitation is a *regional* air pollution problem, not widespread enough to be a *global* problem, but spreading beyond *local* areas. (There have been some unfortunate cases where localized release of acid, usually as sulfur dioxide from metal ore smelting operations have affected local areas, often devastating vegetation within several kilometers of the source

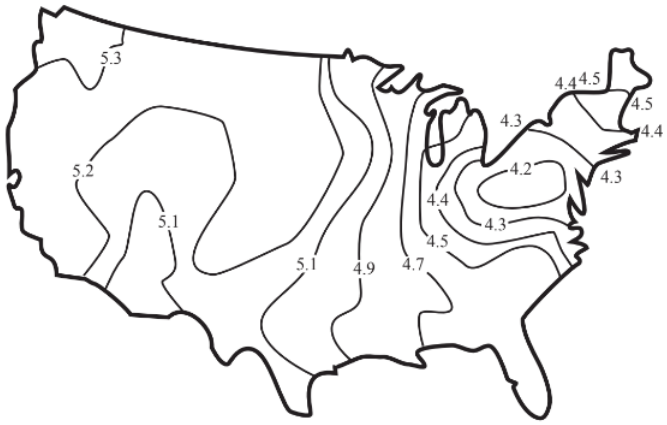


Figure 10.7. Lines showing precipitation with the same pH (isopleths) in the 48 contiguous United States. This is a hypothetical plot. The actual distribution of acidic precipitation varies constantly with time and climatic conditions, but is generally similar to the plot shown with lower values concentrating in the northeastern states.

Transport processes that move atmospheric acids and their precursor acid gases from their sources to downwind areas are very important in determining areas affected by acid rain. The northeastern U.S. and southeastern Canada are affected by acid originating from stack gas emissions carried by prevailing southwesterly winds from Missouri, Illinois, Kentucky and other regions to the southwest. Southern Norway, Sweden, and Finland receive acid precipitation originating farther south in Europe.

Numerous adverse effects have been reported as the result of acidic precipitation. These can be divided into the following major categories:

• Direct effects upon the atmosphere manifested by reduced and distorted visibility. These effects are due to the presence of sulfuric acid droplets and solutions or solid particles of acidic salts, such as NH_4HSO_4.

• Phytotoxicity (toxicity to plants) and destruction of sensitive forests. These effects can be direct, resulting from exposure of plant leaves and roots to acidic precipitation and to acid-forming gases, particularly SO_2 and NO_2. They can also be indirect, primarily by the liberation of phytotoxic Al^{3+} ion by the action of acidic rainfall on soil.

• Direct effects on humans and other animals. These are usually respiratory effects, and asthmatics are especially vulnerable.

• Effects upon plants and fish (especially fish fingerlings) in acidified lake water where the lake is not in contact with minerals, particularly $CaCO_3$, capable of neutralizing acid.

• Damage to materials. Stone (especially acid-soluble limestone and marble) and metal used in building can be corroded and etched by acidic precipitation. Electrical equipment, particularly relay contacts and springs can be corroded by acidic precipitation.

• Some measures can be taken to mitigate the effects of acid rain, although these are very limited once the pollutant has formed. Some success has been achieved with treating acidified lakes with pulverized limestone to neutralize acid. Corrosion-resistant materials can be used in applications where exposure to acid rain is likely. Protective coatings, such as corrosion-resisting paint primers on metals, can be applied to materials likely to be exposed to acidic precipitation. But the best protection is to prevent formation and release of SO_2 and NO_x gases leading to acid rain formation by measures described in the preceding section.

This page titled 2.5.10.8: Acid Rain is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.9: Acid Rain** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.9: Miscellaneous Gases in the Atmosphere

There are several inorganic gases other than oxides that can be significant atmospheric he most common of these is ammonia, NH_3. In addition to industrial pollution, such as from heating coal to make coke for steel making, ammonia can be added to the atmosphere by bacterial sources, from sewage treatment, and from the decay of animal wastes. Accidental releases can occur from liquid anhydrous ammonia used as an agricultural nitrogen fertilizer.

Ammonia is strongly attracted to water, so it is normally present in the atmosphere in water droplets. It is the only significant gaseous base in the atmosphere, so that it reacts with atmospheric acids to produce corrosive ammonium salts as shown by the following reactions:

$$NH_3 + N_2SO_4 \rightarrow NH_4HSO_4$$

$$NH_3 + HNO_3 \rightarrow NH_4NO_3$$

Gaseous chlorine, fluorine, and volatile fluorides are uncommon air pollutants, but very serious where they occur. Elemental chlorine, Cl_2, is widely produced and distributed as a water disinfectant, bleach, and industrial chemical. It is very reactive and so toxic that it was the first poisonous gas used as a military poison in World War I. Most toxic exposures of chlorine occur as the result of transportation accidents leading to its release.

Hydrogen chloride, HCl, can get into the atmosphere from accidental releases of the gas, from reaction with atmospheric water of reactive chlorine-containing chemicals, of which one of the most common is $SiCl_4$,

$$SiCl_4 + 2\,H_2O \rightarrow 2\,SiO_2 + 4\,HCl$$

and from the combustion of chlorine-containing polyvinylchloride (PVC) plastic. The strong affinity of HCl gas for water means that it exists as droplets of hydrochloric acid in the atmosphere. Atmospheric HCl is very irritating to mucous membrane tissue and damaging to corrodable materials.

Elemental fluorine (F_2) and hydrogen fluoride are both highly toxic. Fortunately, occurrences of these gases in the atmosphere are very rare. Gaseous silicon tetrafluoride, SiF_4, can be released during steel making and some metal smelting processes when fluorspar (CaF_2) reacts with sand (SiO_2)

$$2\,CaF_2 + 3\,SiO_2 \rightarrow 2\,CaSiO_3 + SiF_4$$

Sulfur hexafluoride, SF_6, is an astoundingly unreactive gaseous compound used to blanket and degas molten aluminum and magnesium and in gas-insulated electrical equipment. It lasts essentially forever in the atmosphere. As noted in Section 10.6, the greatest concern with its release is that it is a powerful greenhouse warming gas with an effect per molecule about 24,000 times that of carbon dioxide.

Hydrogen sulfide, H_2S, enters the atmosphere from a number of natural sources including geothermal sources, the microbial decay of organic sulfur compounds and the microbial conversion of sulfate, SO_4^{2-}, to H_2S when sulfate acts as an oxidizing agent in the absence of O2. Wood pulping processes can release hydrogen sulfide. Hydrogen sulfide may occur as a contaminant of petroleum and natural gas, and these sources are the most common source of poisoning by H_2S, which has about the same toxicity as hydrogen cyanide. A tragic incident of hydrogen sulfide poisoning occurred in Poza Rica, Mexico, in 1950 as the result of a process to recover H_2S from natural gas. Incredibly, the hydrogen sulfide byproduct was burned in a flare to produce sulfur dioxide. The flare became extinguished at night so that toxic hydrogen sulfide spread throughout the vicinity, killing 22 people and hospitalizing over 300. A 2003 blowout in a natural gas field in southwestern China released hydrogen sulfide that killed almost 200 people in the surrounding area. As an emergency measure the escaping gas was set on fire producing sulfur dioxide, SO_2, still a toxic material, but much less deadly than hydrogen sulfide. Atmospheric H_2S is phytotoxic, destroying immature plant tissue and reducing plant growth. It also affects some kinds of materials, forming a black coating of copper sulfide, CuS, on copper roofing. This coating weathers to a rather attractive green layer (patina) of basic copper sulfate, $CuSO_4 \cdot 3Cu(OH)_2$, which protects the copper from further attack. Hydrogen sulfide in the atmosphere becomes oxidized to SO_2. This process is especially rapid under oxidizing atmospheric conditions such as those in which photochemical smog is formed.

Carbonyl sulfide, COS, is another inorganic sulfur gas that can be detected in the atmosphere, though it is usually at very low levels. A related compound, carbon disulfide, CS_2, also occurs in the atmosphere

This page titled 2.5.10.9: Miscellaneous Gases in the Atmosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.10: Miscellaneous Gases in the Atmosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.10: Photochemical Smog

One of the most common urban air pollution problems is the production of **photochemical smog**. This condition occurs in dry, stagnant air masses, usually stabilized by a temperature inversion (see Figure 8.1), that are subjected to intense sunlight. A smoggy atmosphere contains ozone, O_3, organic oxidants, nitrogen oxides, aldehydes, and other noxious species. In latter stages of smog formation visibility in the atmosphere is lowered by the presence of a haze of fine particles formed by the oxidation of organic compounds in smog.

The chemical ingredients of smog are nitrogen oxides and organic compounds, both released from the automobile, as well as from other sources. The driving energy force behind smog formation is electromagnetic radiation with a wavelength at around 400 nm or less, in the ultraviolet region, just shorter than the lower limit for visible light. Energy absorbed by a molecule from this radiation can result in the formation of active species, thus initiating *photochemical reactions.*

Although methane, CH_4, is one of the least active hydrocarbons in terms of forming smog, it will be used here to show the smog formation process because it is the simplest hydrocarbon molecule. Smog is produced in a series of chain reactions. The first of these occurs when a photon of electromagnetic radiation with a wavelength less than 398 nm is absorbed by a molecule of nitrogen dioxide.

$$NO_2 h\nu \rightarrow NO + O$$

to produce an oxygen atom, O. The oxygen atom is a very reactive species that can abstract a hydrogen atom from methane,

$$CH_4 + O \rightarrow HO \cdot + H_3C \cdot$$

to produce a methyl radical, H3C•, and a **hydroxyl radical**, HO•. In these formulas, the dot shows a single unpaired electron. A chemical species with such a single electron is a **free radical**. The hydroxyl radical is especially important in the formation of smog and in a wide variety of other kinds of photochemical reactions. The methyl radical can react with an oxygen molecule,

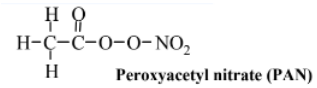
$$H_3C \cdot + O_2 \rightarrow H_3COO \cdot$$

to produce a methylperoxyl radical, H3COO•. This is a strongly oxidizing, reactive species. One of the very important reactions of peroxyl radicals is their reaction with NO, produced in the photochemical dissociation of NO_2 (see Reaction 8.10.1 above),

$$NO + H_3COO \cdot \rightarrow NO_2 + H_3COO \cdot$$

To regenerate NO_2, which can undergo photodissociation, re-initiating the series of chain reactions by which smog is formed. Literally hundreds of other reactions can occur, leading eventually to oxidized organic matter that produces the small particulate matter characteristic of smog.

As the process of smog formation occurs, numerous noxious intermediates are generated. One of the main ones of these is ozone, O_3, and it is the single species most characteristic of smog. Whereas ozone is an essential species in the stratosphere, where it filters out undesirable ultraviolet radiation, it is a toxic species in the troposphere that is bad for both animals and plants. Another class of materials formed with smog consists of oxygen-rich organic compounds containing nitrogen of which peroxyacetyl nitrate, PAN,



Peroxyacetyl nitrate (PAN)

is the most common example. This compound and ones similar to it are potent oxidizers and highly irritating to eyes and mucous membranes of the respiratory tract. Also associated with smog are aldehydes, which are irritants to eyes and the respiratory tract. The simplest aldehyde, and one commonly found in smoggy atmospheres, is formaldehyde:



Formaldehyde

Harmful Effects of Smog

Smog adversely affects human health and comfort, plants, materials, and atmospheric quality. Each of these aspects is addressed briefly here. Ozone is the smog constituent that is generally regarded as being most harmful to humans, plants, and materials,

although other oxidants and some of the noxious organic materials, such as aldehydes, are harmful as well. People exposed to 0.15 parts per million of ozone in air experience irritation to the respiratory mucous tissues accompanied by coughing, wheezing, and bronchial constriction. These effects may be especially pronounced for individuals undergoing vigorous exercise because of the large amounts of air that they inhale. On smoggy days, air pollution alerts may advise against exercise and outdoor activities. Because of these effects, the U.S. Environmental Protection Agency recommends an 8-hour standard limit for ozone of 0.075 ppm and is considering further lowering the standard. In a smoggy atmosphere, the adverse effects of ozone are aggravated by exposure to other oxidants and aldehydes.

Plants are harmed by exposure to nitrogen oxides, ozone, and peroxyacetyl nitrate (PAN, see above), all oxidants present in a smoggy atmosphere. PAN is the most harmful of these constituents, damaging younger plant leaves, especially. Ozone exposure causes formation of yellow spots on leaves, a condition called chlorotic stippling. Some plant species, including sword-leaf lettuce, black nightshade, quickweed, and double-fortune tomato, are extremely susceptible to damage by oxidant species in smog and are used as bioindicators of the presence of smog. Costs of crop and orchard damage by smog run into millions of dollars per year in areas prone to this kind of air pollution, such as southern California.

Materials that are adversely affected by smog are generally those that are attacked by oxidants. The best example of such a material is rubber, especially natural rubber, which is attacked by ozone. Indeed, the hardening and cracking of natural rubber has been used as a test for atmospheric ozone.

Visibility-reducing atmospheric aerosol particles are the most common manifestation of the harm done to atmospheric quality by smog. The smog-forming process occurs by the oxidation of organic materials in the atmosphere, and carbon-containing organic materials are the most common constituents of the aerosol particles in an atmosphere afflicted by smog. Conifer trees(pine and cypress) and citrus trees are major contributors to the organic hydrocarbons that are precursors to organic particle formation in smog.

Preventing Smog with Green Chemistry

Smog is basically a chemical problem, which would indicate that it should be amenable to chemical solutions. Indeed, the practice of green chemistry and the application of the principles of industrial ecology can help to reduce smog. This is due in large part to the fact that a basic premise of green chemistry is to avoid the generation and release of chemical species with the potential to harm the environment. The best way to prevent smog formation is to avoid the release of nitrogen oxides and organic vapors that enable smog to form. At an even more fundamental level, measures can be taken to avoid the use of technologies likely to release such substances, for example, by using alternatives to polluting automobiles for transportation.

The evolution of automotive pollution control devices to reduce smog provides an example of how green chemistry can be used to reduce pollution. The first measures taken to reduce hydrocarbon and nitrogen oxide emissions from automobiles were very much command-and-control and “end-of-pipe” measures. These primitive measures implemented in the early 1970s did reduce emissions, but with a steep penalty in fuel consumption and in driving performance of vehicles. However, over the last three decades, the internal combustion automobile engine has evolved into a highly sophisticated computer-controlled machine that generally performs well, emits few air pollutants, and is highly efficient. (And it would be much more efficient if those drivers who feel that they must drive “sport utility” behemoths would switch to vehicles of a more sensible size.) This change has required an integrated approach involving reformulation of gasoline. The first major change was elimination from gasoline of tetraethyllead, an organometallic compound that poisoned automotive exhaust catalysts (and certainly was not good for people). Gasoline was also reformulated to eliminate excessively volatile hydrocarbons and unsaturated hydrocarbons (those with double bonds between carbon atoms) that are especially reactive in forming photochemical smog.

An even more drastic approach to eliminating smog-forming emissions is the use of electric automobiles that do not burn gasoline. These vehicles certainly do not pollute as they are being driven, but they suffer from the probably unsolvable problem of a very limited range between charges and the need for relatively heavy batteries. However, hybrid automobiles using a small gasoline or diesel engine that provides electricity to drive electric motors propelling the automobile and to recharge relatively smaller batteries can largely remedy the emission and fuel economy problems with automobiles. The internal combustion engine on these vehicles runs only as it is needed to provide power and, in so doing, can run at a relatively uniform speed that provides maximum economy with minimum emissions.

Another approach that is being used on vehicles as large as buses that have convenient and frequent access to refueling stations is the use of fuel cells that can generate electricity directly from the catalytic combination of elemental hydrogen and oxygen, producing only harmless water as a product (see Chapter 16). There are also catalytic process that can generate hydrogen from liquid fuels, such as methanol, so that vehicles carrying such a fuel can be powered by electricity generated in fuel cells.

Green chemistry can be applied to devices and processes other than automobiles to reduce smog-forming emissions. This is especially true in the area of organic solvents used for parts cleaning and other industrial operations, vapors of which are often released to the atmosphere. The substitution of water with proper additives or even the use of supercritical carbon dioxide fluid can eliminate such emissions.

This page titled 2.5.10.10: Photochemical Smog is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.11: Photochemical Smog** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.11: Natural Capital of the Atmosphere

As discussed in Section 1.4, Earth's natural capital is its ability to provide materials, protection, and conditions conducive to life including Earth's resources and its ecosystems. A large fraction of Earth's natural capital is in the atmosphere and includes materials, waste assimilative capacity and esthetics, largely determining the degree to which our surroundings are pleasant and conducive to our existence. The atmosphere's natural capital is discussed in this section.

A huge part of the atmosphere's natural capital is its ability to absorb and protect organisms from destructive ultraviolet and other short-wavelength cosmic and solar electromagnetic radiation which otherwise would make life on Earth impossible. The absorption of longer wavelength infrared radiation by which incoming solar energy is sent back into space leads to the second major protective function of the atmosphere, its ability to maintain surface temperature at a level at which life can thrive (see Figure 10.3).

The atmosphere is a source of essential raw materials, both for organisms and for industrial use and has major applications in the practice of green chemistry. Plants that provide the foundation of food chains within which all organisms thrive extract the carbon dioxide that they use to build biomass from the atmosphere. Animals and other organisms that perform oxic respiration obtain the molecular O_2 they require from the atmosphere. The refractory N_2 in the atmosphere is converted to biomass and protein nitrogen by bacteria growing in soil and water.

Humans also extract gases from the atmosphere for use in the anthrosphere. Adsorption, permeable membrane, and liquified air distillation processes are used to isolate nitrogen, oxygen, argon, and neon from air for use in the anthrosphere. Nitrogen extracted from air is first converted to ammonia, NH_3, then to industrial chemicals, fertilizers, and explosives. Boiling at a frigid-196°C, pure liquid nitrogen is the most widely used cryogenic liquid. Among its many uses are to preserve viable human embryos for embryo implantation to produce "test tube babies." As a generally unreactive gas, pure nitrogen is used as an inert atmosphere that prevents fires and other chemical reactions. Normally extracted from air along with pure nitrogen, pure oxygen has many industrial applications, such as in steel making, and is used for breathing by people with respiratory difficulties. Noble gas argon from the atmosphere is totally chemically inert and is used industrially, such as in specialized welding processes.

Green Oxygen and Nitrogen from the Air

Elemental oxygen and nitrogen are important commercial products extracted from air. Most commonly this is done by distilling cold liquid air, a process that can also produce noble gas neon, argon, and krypton, if desired. The initial step in air distillation is to compress air to about 7 times atmospheric pressure and cool it to remove water vapor and carbon dioxide. Further compression and cooling yields a liquid air product that can then be fractionally distilled to give relatively pure oxygen, nitrogen, and other gases. These can be stored as cold liquids or as the compressed gases.

Essentially pure oxygen has a number of applications, such as for breathing by people with pulmonary insufficiencies. Huge amounts are consumed in steel making. Pure nitrogen is used to provide inert atmospheres free of oxygen. Large quantities of liquid nitrogen are used in the science of **cryogenics** involving very low temperatures.

Emergency Oxygen

Emergency oxygen is required on aircraft that fly at high altitudes. The containers required to transport pure oxygen are too heavy to put on aircraft, so emergency oxygen is generated by a chemical process using a **chlorate candle**. This device contains sodium chlorate, $NaClO_3$, which decomposes when heated to generate oxygen gas:

$$2\,NaClO_3 \rightarrow 2\,NaCl + 3\,O_2$$

Some of the oxygen generated reacts with a fuel, commonly elemental iron, mixed with the sodium chlorate,

$$4\,Fe + 3\,O_2 \rightarrow 2\,Fe_2O_3$$

a heat-generating reaction that provides heat for the decomposition of the sodium chlorate. Chlorate candles can be stored for many years before being activated and still perform well. They are generally safe. However, chlorate candles improperly shipped in the baggage compartment of a ValuJet DC-9 aircraft caused an uncontrollable fire that brought the aircraft down in the Florida Everglades with the loss of all aboard in 1997.

As illustrated in Figure 8.1, the atmosphere is the conduit by which water is evaporated from oceans and carried over land where it falls as precipitation. This ability of the atmosphere is an important component of its natural capital and atmospheric conditions largely determine the quantity, quality, and distribution of water through the hydrologic cycle. Because of variations in atmospheric conditions, the distribution of rainfall is irregular, with excess in some locations and times and deficiencies in others. Hot drought conditions that cause great hardship and even starvation, especially in parts of Africa, are the result of climate conditions in the atmosphere. Sulfur dioxide and nitrogen oxides emitted to the atmosphere as air pollutants produce sulfuric acid and nitric acid, respectively, polluting the hydrosphere with strong acids, killing fish fingerlings and harming vegetation.

Its ability to assimilate and process materials is an important part of the atmosphere's natural capital and a crucial component of nature's natural cycles. Transpiration of water from plant leaves is an important route for conveying water from soil to the atmosphere. Oxic respiration by humans and other organisms discharges carbon dioxide to the atmosphere as do forest fires and anthrospheric combustion processes. Photosynthetically produced elemental oxygen enters the atmosphere. Pollen and small particles such as smoke or fumes produced by anthrospheric processes enter the atmosphere and are washed out with rain or deposited on Earth's surface. Hydrocarbons and nitrogen oxides from combustion are eventually purged from the atmosphere, often with the intermediate formation of oxidants, aldehydes, and particles characteristic of photochemical smog pollution.

The atmosphere's contribution to esthetics is a major facet of its natural capital. Clear, clean air free of visibility-obscuring particles, acidic gases, and ozone that hinders breathing and irritates eyes has genuine value including its contribution to good health. Whereas the water that humans use can be purified from muddy, even polluted sources, air used for breathing must usually be taken as it comes. Humid, foggy air contaminated by acidic constituents and particles is unpleasant and even unhealthy to breathe as is air heated and dried to uncomfortable levels by greenhouse gas emissions.

This page titled 2.5.10.11: Natural Capital of the Atmosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.12: Natural Capital of the Atmosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.12: Energy and Mass Transfer in the Atmosphere

The flux of energy reaching Earth's atmosphere from the sun as sunlight is 1,340 watts/m2.This means that a square meter of area perpendicular to incoming sun rays above Earth'satmosphere is receiving solar energy at a rate sufficient to power 13 100-watt light bulbs plus a 40-watt bulb, enough to power an electric iron or a hair dryer set on high! This is an enormous amount of energy. As shown in Figure 10.3, some of the incoming energy reaches Earth's surface, some is absorbed in the atmosphere, warming it, and some is scattered back to space. The energy that comes in primarily as light at a maximum intensity of 500 nanometers in the visible region must go out, which it does as infrared radiation (with maximum intensity at about 10 micrometers(μm), primarily between 2μm and 40μm). Water molecules, carbon dioxide, methane, and other minor species in the atmosphere absorb some of the outbound infrared, which eventually is all radiated to space. This temporary absorption of infrared radiation warms the atmosphere — a greenhouse effect.

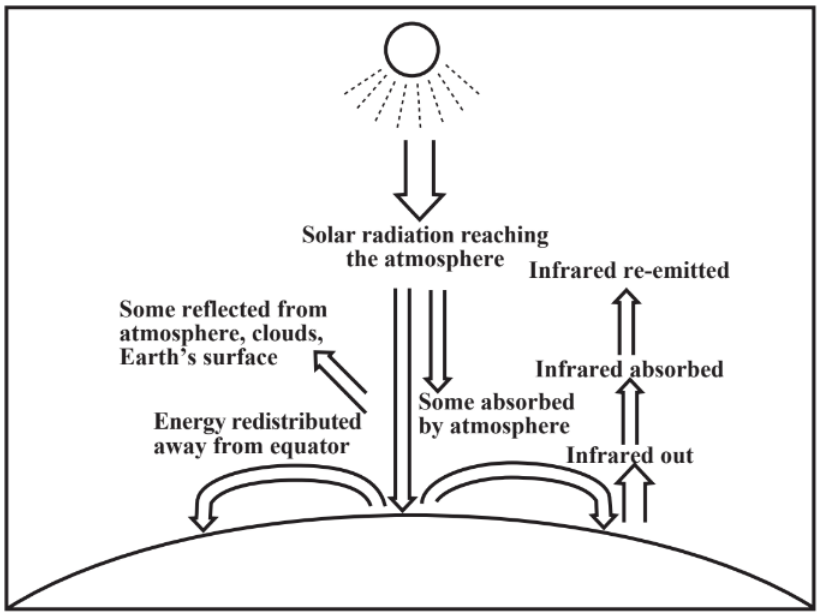


Figure 10.3. Some of the solar energy reaching the top of the atmosphere penetrates to Earth's surface, some is absorbed in the atmosphere, which warms it, and some is scattered by the atmosphere and from Earth's surface. Outgoing energy is in the infrared region. Some of it is temporarily absorbed by the atmosphere before being radiated to space, causing a warming (greenhouse effect). The equatorial regions receive the most energy, part of which is redistributed by warm air masses and latent heat in water vapor away from the equator.

The fraction of electromagnetic radiation from the sun that is reflected by Earth's surface varies with the nature of the surface. The percentage reflected is very important because it determines how effective incoming radiation is in warming the surface and it is expressed as **albedo**. Freshly plowed black topsoil has a very low albedo of only about 2.5%. In contrast, the albedo of a covering of fresh snow is about 90%. The anthrosphere affects albedo. One of the ways that this is done is in cultivating land, turning over relatively high albedo grass and covering it with exposed black soil that absorbs light energy very strongly. Another way is covering of large areas with asphalt paving, which reflects sunlight poorly.

The maintenance of Earth's heat balance to keep temperatures within limits conducive to life is very complex and not well understood. Geological records show that in times past, Earth was sometimes relatively warm and that at other times there were ice ages in which much of Earth'ssurface was covered by ice a kilometer or two thick. The differences in average Earth temperature between these extremes and the relatively temperate climate conditions that we now enjoy were only a matter of a few degrees. It is also known that massive volcanic eruptions and almost certainly hits by large asteroids have caused cooling of the atmosphere that has lasted for a year or more. As addressed later in this chapter, there is now concern that anthropogenic gas emissions, particularly of carbon dioxide from fossil fuel combustion, may be having a warming effect upon the atmosphere.

Earth receives solar energy most directly at the equator, so equatorial regions are warmer than regions farther north and south. A significant fraction of this energy moves away from the equator. This is largely done by **convection** in which heat is carried by masses of air. Such heat can be in the form of **sensibleheat** from the kinetic energy of rapidly moving air molecules (the faster their

average velocities, the higher the temperature). Heat can also be carried as **latentheat** in the form of water vapor. The **heat of vaporization** of water is 2,259 joules per gram (J/g) meaning that 2,259 joules of heat energy are required to evaporate a gram of water without raising its temperature. This is a very high value, meaning that the evaporation of ocean water by solar energy falling on it in warmer regions absorbs an enormous amount of heat to form water vapor. This vapor may be carried elsewhere and condense to form rainfall. The heat energy released raises the temperature of the surrounding atmosphere.

Meteorology

The movement of air masses, cloud formation, and precipitation in the atmosphere are covered by the science of **meteorology**. Meteorologic phenomena have a strong effect upon atmospheric chemistry by processes such as the following:

- Movement of air pollutants from one place to another, such as the movement of air pollutant sulfur dioxide from the U.S. Ohio River Valley to New England and southern Canada, where it forms acid rain.
- Conditions under which stagnant pollutant air masses remain in place so that secondary pollutants, such as photochemical smog, can form.
- Precipitation, which can carry acidic compounds from the atmosphere to Earth's surface in the form of acid rain

Atmospheric chemical processes can influence meterological phenomena. The most obvious example of this is the formation of rain droplets around pollutant particles in the atmosphere.

Weather refers to relatively short term variations in the state of the atmosphere as expressed by temperature, cloud cover, precipitation, relative humidity, atmospheric pressure and wind. Weather is driven by redistribution of energy in the atmosphere. A particularly important aspect of this redistribution is the energy released when precipitation forms. This energy can be enormous because of the high heat of vaporization of water. As an example, heat energy from sunlight and from hot masses of air is converted to latent heat by the evaporation of ocean water off the west coast of Africa. Prevailing winds drive masses of air laden with water vapor westward across the ocean. Rainfall forms, releasing the energy from the latent heat of water and warming the air mass. The hot mass of air that results rises, creating a region of low pressure into which air flows in a circular manner. This can result in the formation of a whirling mass of air in the form of a hurricane that may strike Puerto Rico, Cuba, Florida, or other areas thousands of miles from the area where the water was originally evaporated from the ocean.

A very obvious manifestation of weather consists of very small droplets of liquid water composing **clouds**. These very small droplets may coalesce under the appropriate conditions to form raindrops large enough to fall from the atmosphere. Clouds may absorb infrared radiation from Earth's surface, warming the atmosphere, but they also reflect visible light, which has a cooling effect. Pollutant particles are instrumental in forming clouds. One of the more active kinds of cloud-forming pollutants are atmospheric strong acids, particularly H_2SO_4,

A lack of wind and air currents often occurs under conditions of **temperatureinversion** in which warmer air masses overlay cooler ones (see Figure 10.1). As shown in this figure, topographical features, such as a mountain range that limits horizontal air movement, may make temperature inversion much more effective in trapping polluted masses of air. These conditions occur in the Los Angeles basin noted for photochemical smog formation.

This page titled 2.5.10.12: Energy and Mass Transfer in the Atmosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **10.5: Energy and Mass Transfer in the Atmosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.10.13: Questions and Problems

1. Give two examples each of Earth's natural capital in terms of (1) protective function and (2) raw materials.

2. In what respect is the composition of gases in the troposphere not uniform (which atmospheric constituent varies widely in time and location)?

3. Other than avoiding turbulence due to lower altitude weather, suggest an advantage for commercial aircraft to cruise at a relatively high altitude of around 10 km.

4. Look up the tropopause? Where is it located? What function has it served in keeping Earth in a livable state?

5. What chemical species in the stratosphere is essential for life on Earth?

6. What starts a photochemical reaction? What is it called when a series of photochemical reactions continues?

7. What is a free radical?

8. In what two important respects may very small particles participate in atmospheric chemical processes?

9. In what respect does the radiation by which Earth loses energy differ from that by which it gets energy from the sun?

10. What are two catastrophic events that could cause a sudden cooling of Earth's atmosphere?

11. How is water vapor involved in moving energy through the atmosphere?

12. Distinguish among the terms meteorology, weather, and climate.

13. What do clouds consist of? What must happen before rain falls from clouds?

14. Why is there essentially no atmospheric chemistry involving elemental nitrogen gas in the atmosphere?

15. Cite an atmospheric chemical condition or phenomenon that shows that the O2 molecule iseasier to break apart than the N2molecule.

16. In what respect are elemental nitrogen and oxygen green elements?

17. Give a chemical reaction that produces oxygen that can be used for emergencies.

18. What are two major classes of atmospheric particles based upon how they are produced?

19. In the earlier days of coal utilization, fly ash was not a major problem. What has changed that has resulted in much greater production of fly ash? What modern mode of coal combustion significantly reduces the production of fly ash and acid gases from coal combustion?

20. Suggest why lead has become less of a problem as an atmospheric pollutant in recent years.

21. What is a radioactive element that can get into indoor spaces from underground sources?

22. What is an atmospheric phenomenon caused most prominently by particles 0.1μm–1μm in size? Why are very small particles especially dangerous to breathe?

23. List six means of controlling particle emissions.

24. What is the major health effect of carbon monoxide?

25. What is a serious air pollution phenomenon resulting from an atmospheric reaction of sulfur dioxide?

26. In what form may approximately half of the sulfur in coal be physically separated before combustion?

27. What is a method used to separate sulfur dioxide from furnace stack gas?

28. Name two ways in which green chemistry can be applied to reduce sulfur dioxide emissions.

29. What is an important health effect of nitrogen dioxide? Why is it particularly important in atmospheric chemistry?

30. In 2008/2009 problems arose in newer houses due to toxic drywall. What was the cause of this problem and how does it relate to material covered in this chapter?

31. What are five categories of adverse effects from acid precipitation?

32. Chemically, what is distinctive about ammonia in the atmosphere?

33. What is the historic evidence for the toxicity of elemental chlorine in the atmosphere?

34. What are some sources of atmospheric hydrogen sulfide? Is it a health concern?

35. In what respect is atmospheric carbon dioxide essential to life on Earth? Why may it end up being the “ultimate air pollutant?”

36. What are some of the more harmful effects projected if global warming occurs to a significant extent?

37. What can green chemistry do about global warming?

38. What is a greenhouse gas other than carbon dioxide that is produced by microorganisms?

39. What are the ingredients and conditions leading to the formation of photochemical smog?

40. What substances are found in a smoggy atmosphere?

41. What are some harmful effects of smog?

42. What are some of the ways that green chemistry can help prevent smog?

43. The temperature of a specific number of moles of gas occupying initially 23.0 L was changed from 75°C to -20°C at constant pressure. Recalling the significance of 273 in such calculations, what was the volume of the gas after the temperature change?

44. The pressure on a specific number of moles of gas occupying initially 13.0 L was changed from 1.15 atm to 0.900 atm at constant pressure. What was the volume of the gas after the temperature change?

2.5.10.13: Questions and Problems is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** has no license indicated.

2.5.10.14: Supplementary References

Aguado, Edward, and James E. Burt, *Understanding Weather and Climate*, 4th ed., Pearson Education, Upper Saddle River, NJ, 2007.

Ahrens, C. Donald, *Meteorology Today: An Introduction to Weather, Climate, and the Environment*, 9th ed., Thomson Brooks/Cole, Belmont, CA, 2009.

Ahrens, C. Donald, *Essentials of Meteorology Today: An Invitation to the Atmosphere*, 5th ed., Thomson Brooks/Cole, Belmont, CA, 2008.

Allaby, Michael, *Atmosphere: A Scientific History of Air, Weather, and Climate*, Facts on File, New York, 2009.

Andrews, David G., *An Introduction to Atmospheric Physics*, Cambridge University Press, Cambridge, U.K., 2010.

Austin, Jill, Peter Brimblecombe, William Sturges, Eds., *Air Pollution Science for the 21st Century*, Elsevier Science, New York, 2002.

Barker, John. R., "A Brief Introduction to Atmospheric Chemistry," *Advances Series in Physical Chemistry*,**3**, 1-33 (1995).

Desonie, Dana, *Atmosphere: Air Pollution and its Effects*, Chelsea House Publishers, New York,2007.

Frederick, John E.,*Principles of Atmospheric Science*, Jones and Bartlett, Sudbury, MA, 2008.

Garratt, Richard, *Atmosphere: A Scientific History of Air, Weather, and Climate*, Facts on File, New York, 2009.

Hewitt, C. N., and Andrea Jackson, *Atmospheric Science for Environmental Scientists*, Wiley-Blackwell, Hoboken, NJ, 2009.

Hewitt, Nick, and Andrea Jackson, Eds., *Handbook of Atmospheric Science*, Blackwell Publishing, Malden, MA, 2003.

Hidore, John J., John E. Oliver, Mary Snow, and Richard Snow, *Climatology: An Atmospheric Science*, 3rd ed., Prentice Hall, Upper Saddle River, NJ, 2009.

Hobbs, Peter V., *Introduction to Atmospheric Chemistry*, Cambridge University Press, New York,2000.

Jacob, Daniel J., *Introduction to Atmospheric Chemistry*, Princeton University Press, Princeton, NJ, 1999.

Spellman, Frank R., *The Science of Air: Concepts and Applications*, 2nd ed., Taylor & Francis, Boca Raton, FL, 2009.

Vallero, Daniel, *Fundamentals of Air Pollution*, 4th ed., Elsevier, Amsterdam, 2007.

Wallace, John M., *Atmospheric Science: An Introductory Survey*, 2nd ed., Elsevier Amsterdam,2006

2.5.10.14: Supplementary References is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Supplementary References** has no license indicated.

SECTION OVERVIEW

2.5.11: The Geosphere and a Green Earth

“The 2010 explosion, fire, and subsequent massive leakage of crude oil at the site of the Deepwater Horizon oil well in the Gulf of Mexico illustrates the hazards of probing deep into the geosphere without proper safeguards against the kinds of accidents that can occur at great depths where high temperatures and extreme pressures present unique challenges to technology.”

2.5.11.1: Probing Deep into the Geosphere - A Well Too Far?

2.5.11.2: The Nature of the Geosphere

2.5.11.3: The Geosphere as a Source of Natural Capital

2.5.11.4: Environmental Hazards of the Geosphere

2.5.11.5: Water on and in the Geosphere

2.5.11.6: Anthrospheric Influences on the Geosphere

2.5.11.7: The Geosphere as a Waste Repository

2.5.11.8: Have You Thanked a Clod Today?

2.5.11.9: Plant Nutrients and Fertilizers in Soil

2.5.11.10: Soil and Plants Related To Wastes and Pollutants

2.5.11.11: Soil Loss - Desertification and Deforestation

2.5.11.12: Literature Cited and Supplementary References

2.5.11.13: Questions and Problems

This page titled 2.5.11: The Geosphere and a Green Earth is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.11.1: Probing Deep into the Geosphere - A Well Too Far?

In 2009 the British petroleum exploration and development company, BP, reached a record depth for an oil well of 10,685 meters (35055 feet, 6.6 miles) below the ocean floor in the Gulf of Mexico not counting the 1,220 meters of water above the wellhead. This well is part of the Tiber Prospect oil field estimated to contain between 4 billion and 6 billion barrels of crude oil equivalent (including natural gas). This well illustrates the ability that humans have acquired to probe far into the geosphere where conditions are extreme with pressures that may reach 1,200 times atmospheric pressure and temperatures may exceed 135°C. At these depths and under these conditions it is indeed a different world than the one encountered by more conventional petroleum exploration.

In early 2010 BP was in the process of completing another deep well in the Gulf of Mexico about 66 km off the coast of Louisiana in 1500 meters of water using the Deepwater Horizon semi-submersible mobile offshore drilling unit. This marvel of modern well drilling technology was a massive floating dynamically-positioned unit kept in place precisely by the computerized operation of propulsion units. At 9:45 p.m. on April 20, 2010, highly pressurized methane gas overcame the containment devices on the well, burst through the top of the drill column, and spread over the drill rig causing a massive explosion and huge fire. Although most of the personnel on the rig were rescued, 11 workers were killed and their bodies were never found. After 36 hours of uncontrollable fire, the Deepwater Horizon rig sunk on the morning of April 22, 2010. Despite efforts to contain the flow, oil pressurized by natural gas continued to pour from the wellhead, spreading across large areas of the Gulf of Mexico resulting in arguably the most catastrophic environmental disaster of our time. Financial costs of this disaster have been in the billions of dollars (in June 2010 BP set aside a $20 billion fund to pay claims resulting from the oil release) and damage to fisheries and coastal areas from the leaking petroleum have been immeasurable. The question can be raised whether this was “a well too deep,” a probing by humans too far into the depths of the geosphere, at least in consideration of the inadequate protective measures that had been taken?

This page titled 2.5.11.1: Probing Deep into the Geosphere - A Well Too Far? is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.1: Probing Deep into the Geosphere - A Well Too Far?** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.11.2: The Nature of the Geosphere

Defined in Section 8.4 and illustrated in Figure 8.5, the **geosphere** consists of the rocks, minerals, soils, sediments, molten magma, and pressurized gases and liquids beneath Earth's surface. The Deepwater Horizon incident discussed above speaks to human efforts to extract from the geosphere the materials and energy that modern civilization demands and it illustrates the potential environmental hazards associated with exploiting the geosphere's resources. It speaks to the need for understanding the geosphere and treating it with the respect that it demands for the maintenance of the environment under living conditions.

Most of our food is grown on the geosphere and humans extract from it metals, fossil fuels, fertilizers for plants, and a variety of minerals used in construction and for other purposes. Over the years, huge quantities of waste products have been discarded to the geosphere, in some cases very carelessly in a manner that poses threats to humans and other organisms. A thin layer on top of the geosphere — in places only a few centimeters deep — composes topsoil which supports the plant life upon which humans and most other land-dwelling organisms depend for their food.

The geosphere interacts strongly with the other environmental spheres. Streams and rivers flow through channels in the geosphere, lakes and reservoirs occupy cavities on the surface of the geosphere, and groundwater occurs in aquifers underground that are part of the geosphere. Gases are exchanged between the geosphere and the atmosphere, light and infrared radiation transmitted through the atmosphere warm the surface of ground, and it in turn radiates back to the atmosphere the infrared radiation by which Earth loses the energy it absorbs from the sun.

The geosphere is tied to green chemistry in many important respects, including the following

- Plants that provide most food for humans and animals grow on the geosphere.
- Plants growing on the geosphere already provide, and have the potential to provide much more, biomass for use as renewable materials, such as wood, fiber, raw materials, and fuel.
- The geosphere is the source of nonrenewable minerals, ores, fossil fuels, and other materials used by modern industrialized societies.
- Modifications and alterations of the geosphere have profound effects upon the environment.
- Sources of fresh water are stored in lakes and rivers on the surface of the geosphere, move by means of streams, rivers, and canals on the geosphere, and occur in aquifers underground.
- The geosphere is the ultimate sink for disposal of a variety of wastes.

This chapter briefly addresses the nature of the geosphere, and resource utilization from the geosphere. Waste disposal on land or underground are considered in later chapters. Because of the special importance of soil and the plants that grow on it as sources of food and fiber, soil is discussed in some detail.

Physical Nature of the Geosphere

At the center of Earth is an iron-rich inner core, hot enough to be molten under normal pressures, but compressed to a solid by the enormous pressures at such great depths. Surrounding this core is an outer core consisting of molten rock called magma. Earth's solid outer layer consists of the mantle and the **crust**, a layer that is only 5-40 km thick. Only the upper layers of the crust are accessible to humans.

For the most part, the crust consists of **rocks**, which in turn are made up of **minerals** characterized by a definite chemical composition and crystal structure. Only about 25 of the approximately 2000 known minerals compose most rocks. Because most of the crust consists of chemically combined oxygen (49.5%) and silicon (25.7%), the most abundant minerals are **silicates** composed of various silicon oxides, examples of which are quartz, SiO_2, and potassium feldspar, $KAlSi_3O_8$. Other elements in Earth's crust are aluminum (7.4%, commonly occurring as Al_2O_3), iron (4.7% as Fe_3O_4 and other iron oxides), calcium (3.6% in limestone, $CaCO_3$, and dolomite, $CaCO_3{\bullet}MgCO_3$), sodium (2.8%), potassium (2.6%), and magnesium (2.1%). That leaves only 1.6% of the crust to serve as a source of other important mineral substances, including metals other than iron and aluminum, phosphorus required for plant growth, and sulfur widely used in industrial applications.

The rocks that compose Earth's crust participate in the **rockcycle** shown in Figure 11.1. **Igneous rock** is rock that has solidified from molten rock called **magma** that has penetrated to near Earth's surface. Exposed to water, atmospheric oxygen, and various organisms, igneous rock becomes highly altered, reaching a state of greater physical and chemical equilibrium with the atmosphere.

This is a process called **weathering**. Weathering products end up as soil and are carried by water to be deposited as sediments. Sediments that become buried and compressed become **secondary minerals**, among the most abundant of which are **clays**, consisting of silicon and aluminum oxides, produced by the weathering of minerals such as potassium feldspar, $KAlSi_3O_8$. A common clay is **kaolinite**, $Al_2Si_2O_5(OH)_4$.

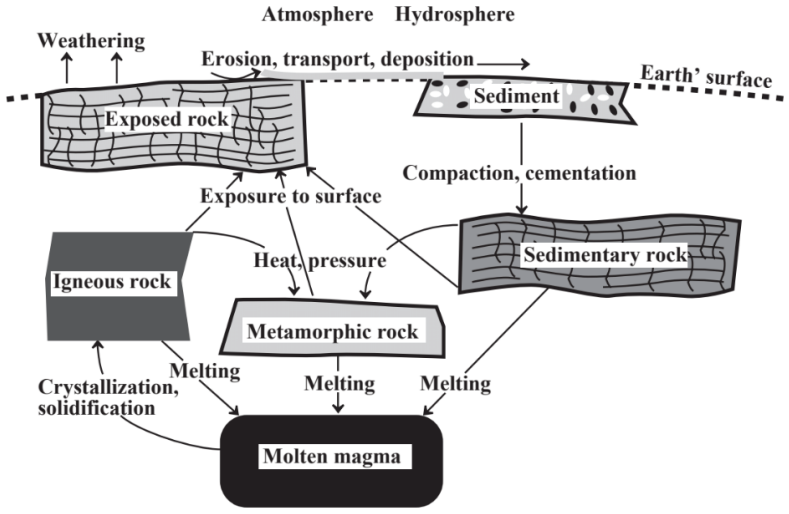


Figure 11.1. The rock cycle. Rock is produced when molten magma crystallizes and solidifies to produce igneous rock. Igneous rock is exposed to the atmosphere and hydrosphere at the surface where weathering and erosion produces sedimentary material that is deposited by the movement of water in sediments. The sediments become buried and subjected to pressure that result in compaction and cementation to produce sedimentary rock. Heat and pressure generate metamorphic rock. At sufficient depths rock of all kinds can be subjected to high temperatures to produce molten rock called magma. Magma that has crystallized and solidified is igneous rock. Igneous rock forced to the surface and exposed to water and air weathers and produces sedimentary material to start the rock cycle over.

Although the earth's crust is very thin compared to Earth's total diameter, there is an even much thinner, fragile and vitally important layer covering the crust —**soil**. Soil is the finely divided mixture of mineral and organic matter upon which plants grow, providing the food that humans and most other animals eat. Productive soil may be only a few centimeters thick, and rarely is more than a few meters in thickness. Soil is uniquely important and of crucial importance in sustainability. Humans are capable of inflicting great damage on soil causing it to become unproductive and in extreme cases resulting in widespread hunger and even starvation.

Geochemistry is the branch of chemistry that deals with rocks and minerals and the chemical interactions of the geosphere with other environmental spheres. The specialized branch of geochemistry relating to environmental influences and interactions of the geosphere is **environmental geochemistry**. Weathering by chemical processes is a particularly important aspect of geochemistry. Almost imperceptible under dry conditions, weathering proceeds at a much more rapid rate in the presence of water. The rate of weathering is also increased by the action of microorganisms, some of which secrete chemical species that attack rock and leach nutrients from it. Particularly important to weathering are **lichens**, which are algae and fungi living together synergistically. The algae utilize solar energy to convert atmospheric carbon dioxide to plant biomass and the fungi utilize the biomass and anchor the organisms to the rock surface and extract nutrients from it.

Human Influences

Human activities have a tremendous influence on the geosphere as evidenced by hills leveled, valleys filled in, and vast areas paved to make freeways, parking lots, and shopping centers. One such influence is on **surface albedo**, defined as the percentage of impinging solar energy reflected back from Earth's surface. The surface albedo of an asphalt paved surface is only about 8%. Amore alarming effect is desertification in which normally productive soil is converted to unproductive desert in areas where rainfall is marginal. This phenomenon is discussed in more detail in Section 11.10.

This page titled 2.5.11.2: The Nature of the Geosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.2: The Nature of the Geosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

https://bio.libretexts.org/@go/page/108873

2.5.11.3: The Geosphere as a Source of Natural Capital

The geosphere is an immense source of natural capital, providing a living environment for most humans, minerals required by modern civilizations, and room for disposal of wastes. One of the greatest concerns with regard to sustainability is the acquisition of essential elements from the geosphere in ways that sustain to the maximum extent possible supplies of these irreplaceable resources. As technology advances priorities for specific geospheric resources change. In recent years, numerous uses have emerged for the **rare earth elements** consisting of the fifteen lanthanides (elements with atomic numbers 58 through 71 in the periodic table shown in Chapter 3) plus scandium and yttrium, transition elements with atomic numbers 21 and 39. The chemical properties of the lanthanides are generally quite similar making their separation difficult and the properties of scandium and yttrium are similar to those of the lanthanides, so they are commonly classified as rare earths.

The uses that have emerged for the rare earths are varied and for different ones include making metal alloys, superconductors, phosphors that glow various colors in light-emitting diodes(LEDs), electrodes, electrolytes, electronic filters, lasers, specialty (colored) glasses, X-ray tubes, mercury vapor lamps, computer memory, oxidizing agents, and reducing agents. Rare earth's are widely used in hybrid automobiles and in wind turbines. Each Toyota Prius hybrid automobile reportedly requires 1 kg of neodymium for its electric motor with terbium and dysprosium added in smaller quantities to preserve magnetic qualities and 10-15 kg of lanthanum for its electric motor.

Lithium has emerged as an important element because of the emergence of high-powered lithium storage batteries capable of storing and releasing large quantities of energy per unit mass of battery. These have largely been used in computers and other electronic devices, but will certainly find growing applications in electric and hybrid automobiles.

With the rather sudden development of new applications for rare earth elements and lithium, questions of supply have become important. China has had a near monopoly on rare earth elements and, with the advancement of high-tech industries in China which may consume available supply, other countries have become alarmed regarding availability. China is also a source of lithium, although Bolivia is the main supplier. Fortunately, rare earths are not very rare and the vast U.S. deposit in Mountain Pass, California, was the largest supplier until the facility was closed due to competition from China where labor costs are very low. Now the mine is being modernized in preparation for production to resume. Rare earth deposits also occur in Canada and even Vietnam.

In June 2010, U.S. military officials and geologists revealed that war-torn Afghanistan was a treasure trove of desired minerals including rare earths with a total value of all mineral resources estimated at around 1 trillion dollars. The most abundant and valuable of these is iron (estimated at $420 billion), copper ($274 billion), niobium ($81 billion), cobalt ($51 billion), gold ($25 billion), molybdenum ($24 billion), and rare earths ($7.4 billion). Other minerals of commercial value in Afghanistan likely include silver, potash, aluminum, graphite, fluorite, phosphorus, lead, zinc, mercury, strontium, sulfur, talc, magnesite, and kaolin clay. There are also believed to be lithium deposits in dry lake beds of Afghanistan's eastern province of Ghazni. The lithium deposits may in fact be equal to those of Bolivia, which currently produces most of the lithium used in battery manufacture. Development of these mineral sources has the potential to help move the economy of the troubled country of Afghanistan from dependence on the opium trade (and U.S. military expenditures) to an economy based upon mineral resources.

This page titled 2.5.11.3: The Geosphere as a Source of Natural Capital is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.3: The Geosphere as a Source of Natural Capital** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.11.4: Environmental Hazards of the Geosphere

Having lain dormant for almost two centuries, the Eyjafjallajokull volcano, one of Iceland'slargest, began to ooze lava on March 20, 2010, visible as a red glow above the huge glacier covering the volcano. Initially, the eruption was nothing more than an interesting tourist attraction and the volcano appeared to revert to its normal state after a few days. However, on April 14 an enormous explosion sent volcanic ash as high as 11,000 meters followed by days in which the volcano spewed ash high into the atmosphere. This presented a significant problem for commercial aviation as the plume of volcanic ash spread eastward across the British Isles and northern and central Europe because volcanic ash can damage jet engines and even cause them to stop running. (In 1982 all four engines of a British Airways 747 stopped when it inadvertently flew into an ash cloud from Indonesia's Mount Galunggung and for several terrifying minutes what suddenly became the world's largest glider descended from 11,000 meters to 4100 meters before the engines restarted enabling an emergency landing in Jakarta.) The result of the Eyjafjallajokull eruption was that within two days most of Europe's major airports were closed canceling thousands of flights. Because of ripple effects across the world this incident became the worst peacetime travel disruption in history stranding millions of travelers, many with diminished financial resources from limited travel budgets. The result was a period of many days of travel chaos as flight bookings were rescheduled to eventually get travelers to their destinations. Airlines estimated financial losses of about $1.7 billion resulting from the cancellation of more than 100,000 flights.

The Iceland volcano eruption has been a reminder of the awesome, potentially destructive forces that reside in the geosphere and of the vulnerability of modern civilization to them. Along with earthquakes, volcanoes are geospheric phenomena that are beyond the power of humans to prevent or even accurately predict. Even in these cases, however, human activities can significantly influence the degree of damage done. As examples, structures constructed on poorly consolidated fill dirt are much more susceptible to earthquake damage than are those attached firmly to bedrock, and the construction of dwellings in areas known to be subject to periodic volcanic eruptions simply means that unstoppable lava flows and other volcanic effects will be much more damaging when they occur. Other, less spectacular, but very destructive geospheric phenomena can be greatly aggravated by human activities. Destructive and sometimes life-threatening landslides, for example, often result from human alteration of surface soil and vegetation.

Earthquakes

Earthquakes consist of violent horizontal and vertical movement of Earth's surface resulting from relative movements of tectonic plates. Plates move along fault lines. Huge masses of rock may be locked relative to each other for as long as centuries, then suddenly move along fault lines. This movement and the elastic rebound of rocks that occurs as a result causes the earth to shake, often violently and with catastrophic damage.

History provides many examples of astoundingly damaging earthquakes. Over 1 million lives(out of a much lower global population than now) were lost by an earthquake in Egypt and Syria in 1201 A.D. The Tangshan, China, earthquake of 1976 killed approximately 650,000. During the latter 1990s and early 2000s, a number of fatalities resulted from earthquakes in Turkey, Greece, Taiwan, Iran, India and China. The May 12, 2008, 7.9 magnitude Wenchuan earthquake in Sichuan Province, China, left 80,000 people dead or missing. Financial costs of earthquakes in highly developed areas are enormous; the 1989 Loma Prieta earthquake in California cost about 7 billion dollars. Phenomena caused by earthquakes can add to their destructiveness. In addition to their direct shaking effects, earthquakes can cause ground to rupture, subside, or rise. **Liquefaction** of poorly consolidated ground, especially where groundwater levels are shallow, occurs when soil particles disturbed by an earthquake separate and behave like a liquid, causing structures to sink and collapse. One of the more terrifying effects of earthquakes are giant ocean waves called **tsunamis** that can reach heights of as much as 30 meters. On December 26, 2004, a huge earthquake off the coast of Sumatra generated a tsunami up to 30 meters high, killing more than 150,000 people in countries around the Indian Ocean.

Earthquakes have defied all efforts to predict them, a fact that makes them all the more frightening. However, earthquake-prone areas, such as southern California, are well known, and loss of life and property can be minimized by taking appropriate measures. Buildings can be constructed to resist the effects of earthquakes using practices that have been known for sometime. For example, some buildings in Niigata, Japan, were constructed to be earthquake-resistant in the 1950s. When a destructive earthquake hit that city in 1964, some buildings tipped over on the liquified soil but remained structurally intact! (Current practice calls for the construction of more flexible structures designed to dissipate the energy imparted to them by an earthquake.) The construction of

buildings, roadways, railroads, and other structures to withstand the destructive effects of earthquakes provides an excellent example of designing the anthrosphere in a manner that is as compatible as possible with the geosphere and the natural hazards it poses.

Although humans can do nothing to prevent earthquakes, there is some evidence that anthrospheric activities have helped cause them. Some seismologists have suggested that the pressure of water from newly constructed reservoirs in China provided lubrication that enabled earth movement. At least one experiment in injecting water into hot rock formations to produce steam for power had to be stopped because it caused a number of very small quakes detected by sensitive instruments.

Volcanoes

A volcano results due to the presence of liquid rock magma near the surface. In addition to liquid rock lava at temperatures ranging from 500°C to 1400°C that flows from volcanoes, these often very destructive phenomena are manifested by discharges of gases, steam, ash, and particles. Volcanic disasters have always plagued humankind. The 79 A.D. eruption of Mount Vesuvius in ancient Rome buried the city of Pompei in ash, preserving a snapshot of life in Rome at that time. The astoundingly massive eruption of Indonesia's Tambora volcano in Indonesia in 1815 was caused when water infiltrated the hot magma beneath the volcano resulting in an explosion equivalent to 100 million tons of TNT explosive and blasting an estimated 30 cubic kilometers of solid material into the atmosphere. The May 18, 1980, Mount St. Helens eruption in Washington State blew about 1 cubic kilometer of material into the atmosphere, killed 62 people, and caused about $1 billion in damage.

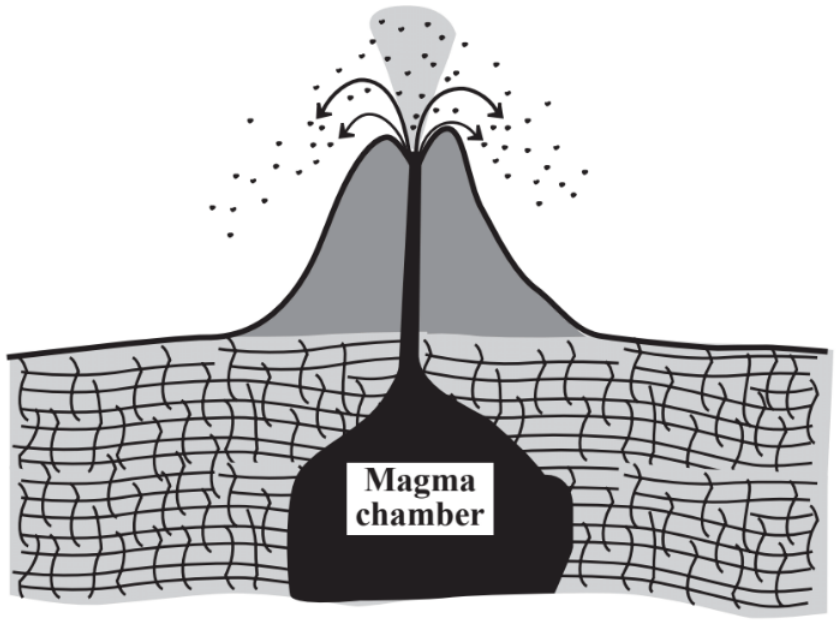


Figure 11.2. A volcano in a classic cone shape produced when molten lava and ash are ejected from a magma chamber underground.

In addition to their immediate effects upon surrounding areas, volcanoes can affect the atmosphere and climate. The Tambora volcano blasted enough particulate matter into the atmosphere to cause a very pronounced cooling effect. The following "year without a summer" caused global crop failures and starvation, and perceptible global cooling was observed for the next 10 years. Huge quantities of water vapor, dense carbon dioxide gas, carbon monoxide, hydrogen sulfide, sulfur dioxide, and hydrogen chloride may be emitted to the atmosphere in volcanic eruptions. People may suffocate in the carbon dioxide or be poisoned by the toxic carbon monoxide and hydrogen sulfide. Hydrogen chloride along with hydrogen sulfide and sulfur dioxide oxidized in the atmosphere to sulfuric acid can contribute to acidic rainfall. Volcanic emissions differ in their atmospheric chemical effects. The 1982 El Chichón eruption in Mexico generated little particulate mineral matter but vast amounts of sulfur oxides that were oxidized to sulfuric acid in the atmosphere. The tiny droplets of sulfuric acid suspended in the atmosphere effectively reflected enough sunlight to cause a perceptible cooling in climate.

Massive, atmospheric-damaging eruptions of volcanoes in recorded history have caused catastrophic crop failures. These will happen again. And since the world as a whole carries little food surplus from year to year, the certainty of food supply disruptions

due to volcanic activity point to the desirability of storing substantial amounts of food for emergency use.

Surface Effects

Though less spectacular than major earthquakes or volcanic eruptions, surface earth movement causes enormous damage and significant loss of life. Furthermore, surface earth movement is often strongly influenced by human activities. Surface phenomena result from the interaction of forces that act to thrust earth upward countered by weathering and erosion processes(see Section 11.2) that tend to bring earth masses down. Both of these phenomena are influenced by the exposure of earth masses to water, oxygen, freeze-thaw cycles, alternate saturation with water and drying, organisms and human influences.

Landslides occur when finely divided (unconsolidated) earthen material slides down a slope.The results can be devastating. The 1970 earthquake-initiated landslide of dirt, mud, and rocks on the slopes of Mt. Huascaran in Peru may have killed 20,000 people. A 1963 landslide on slope surrounding a reservoir held by the Vaiont Dam in Italy suddenly filled the reservoir causing a huge wall of water to overflow the dam killing 2600 people and destroying everything in its path.

Along with weather and climate, human activities can influence the likelihood and destructiveness of landslides. Roads and structures constructed on sloping land can weaken the integrity of earthen material or add mass to it, increasing its tendency to slide. In some cases, strong root structures of trees and brush anchor sloping land in place. However, some plant roots destabilize and add mass to soil, increase the accumulation of water underground, and cause earth to slide. Fortunately, predicting a tendency for landslides to occur is relatively straightforward based upon the nature and slope of geological strata, climate conditions, and observations of evidence of a tendency toward landslides, such as movement of earth and evidence of cracked foundations in buildings built on slopes. In some cases remedial actions may be taken, but more important are the indications that structures should not be built on slide-susceptible slopes.

Less spectacular than landslides is **creep** characterized by a slow, gradual movement of earth. Creep is especially common in areas where the upper layers of earth undergo freeze/thaw cycles. A special challenge is **permafrost** which occurs in northern Scandinavia, Siberia, and Alaska. Permafrost refers to a condition in which ground at a certain depth never thaws, and thawing occurs only on a relatively thin surface layer. Structures built on permafrost may end up on a pool of water-saturated muck resting on a mixture of frozen ice and soil. One of the greater challenges posed by permafrost in recent times has been the construction of the Trans-Alaska pipeline in Alaska on a permafrost surface. Global warming is causing thawing of permafrost in Arctic regions such as parts of Siberia and is resulting in significant structural damage.

Expansive clays that alternately expand and contract when saturated with water, then become dried out, can cause enormous damage to structures, making the construction of basements virtually impossible in some areas. **Sinkholes** occur in areas where rock formations are dissolved by chemical action of water (particularly dissolved carbon dioxide acting on limestone). Earth can fall into a cavity generated by this phenomenon causing huge holes in the ground that can swallow several houses at a time.

This page titled 2.5.11.4: Environmental Hazards of the Geosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.4: Environmental Hazards of the Geosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.11.5: Water on and in the Geosphere

The geosphere is the repository of virtually all the world's fresh water. As shown in Figure11.2, this water may be in underground aquifers as groundwater; on the surface as streams, rivers, lakes and impoundments; or as deposits of ice (glaciers) resting on Earth's surface. Water collected

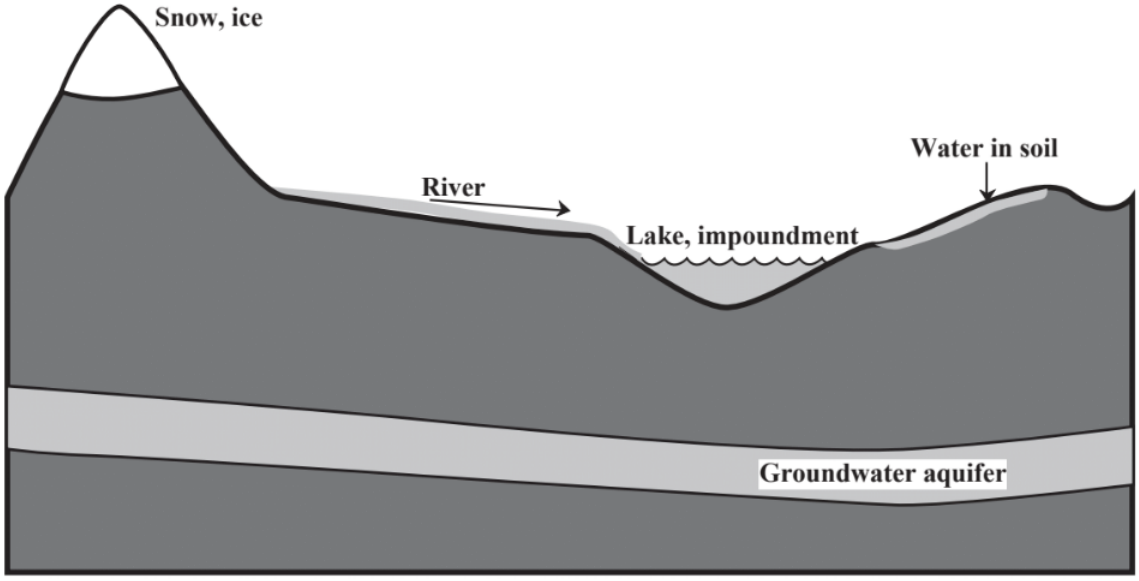


Figure 11.3. Major aspects of water on and in the geosphere

by the geosphere constitutes virtually all Earth's fresh water resources. This water is susceptible to pollution which in extreme cases can render the water sources virtually useless. One of the greater water pollution problems is when water in underground aquifers becomes contaminated with hazardous waste material improperly discarded in the geosphere.

Water commonly moves on the geosphere in **streams** or **rivers** consisting of channels through which water flows. Rivers collect water from drainage basins or watersheds. In order to protect water quality in rivers, pollution and pollution-causing agricultural practices in drainage basins must be avoided. Rivers continually erode the geosphere over which they flow and leave deposits of sediments. Over time, a river will erode earth away and create valleys. An undisturbed river continually cuts curving patterns known as **meanders** in a river valley. The flat area of a valley formed by erosion and sediment deposition in the valley and susceptible to periodic flooding is the river's floodplain.

Floods are the phenomena associated with river flow that are most likely to cause damage to human structures. Despite their destructiveness, floods are normal phenomena by which a river does much of its work of shaping the surface of the geosphere. However, by unwisely building in flood plains, humans have made themselves susceptible to the damaging effects of floods. This was illustrated most tragically by the deadly flood of the city of New Orleans following the 2005 Hurricane Katrina in which many of the areas flooded were built below sea level! Human activities on the geosphere surface can make the effects of floods much worse. For example, flash floods following intense rainfall in urban areas are made much worse by the removal of vegetation from watersheds and its replacement with paving. Concrete and asphalt surfaces do not slow down the flow of water like well-rooted plants do and such surfaces prevent the infiltration of water into the ground.

Attempts to control water flow and flooding provide interesting examples of how humans can interact with their natural environment. Control measures have concentrated on the downstream end on the rivers themselves by construction of levees to confine rivers to their banks, straightening and deepening river channels to increase the velocity and flow of the water in an effort to move it quickly downstream away from the potentially flooded area, and by building dams to contain floodwater until it can be safely released. Such measures can be deceptively successful, sometimes for many decades, until a massive flood overwhelms them. When a contained river carrying vast amounts of water flowing at a high velocity eventually breaks through the levees and dams designed to contain it, the resulting damage can be catastrophic.

An approach to flood control based upon the best practice of sustainability provides a means of minimizing flood damage. Such an approach tends to concentrate more on the upstream end, the watersheds from which water produced by rainfall flows into the river. With the proper kind of vegetation cover, such as forests, and with terraces and small dams designed to temporarily slow the flow of water into the river from the watershed, extremes of high water (flood crests) can be greatly reduced. With regard to protection of dwellings and agricultural land in the river's floodplain, a fundamental question has to be asked whether houses should even be located in these areas and whether the land should be cultivated. In many cases, the answer is no, and the least costly alternative overall is to pay for removal of the structures and conversion of the land back to an uncultivated state, simply allowing the flooding that comes naturally to the river.

This page titled 2.5.11.5: Water on and in the Geosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.5: Water on and in the Geosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.11.6: Anthrospheric Influences on the Geosphere

The urge to “dig in the dirt” and alter Earth surfaces seems to be innate in humans. During recent decades the potential of humans to alter the geosphere has been greatly increased by the development of massive earth-moving equipment. Flooding of rivers caused by human activities was discussed above. Other geospheric disturbances that can be detrimental include landslides on mounds of waste mine tailings, adverse effects resulting from exposure of minerals during mining (production of acid mine water from exposure of pyrite, FeS_2, in coal mining), and filling and destruction of wetlands upon which many forms of wildlife depend for breeding grounds.

Human effects upon the geosphere can be both direct and indirect. Construction of dams and reservoirs, flattening whole mountain tops to get to underground coal seams, and plowing natural prairies to grow crops are obvious direct effects. Indirect effects include pumping so much water from underground aquifers that the ground subsides, or exposing minerals to the atmosphere by strip mining so that the minerals undergo weathering to produce polluted acidic water. In extracting minerals from the earth, it is disturbed and rearranged in ways that can cause almost irreversible damage to the environment. A major objective of the practice of green chemistry and industrial ecology is to minimize these detrimental effects and, to the extent possible, eliminate them entirely.

Many of the effects of human activities on the geosphere have to do with the extraction of resources of various kinds from Earth’s crust. These may range from gravel simply scooped from pits on Earth’s surface to precious metals at such low concentrations that tons of ore must be processed to get a gram or less of the metal. The most straightforward means of obtaining materials from Earth’s crust is surface mining. This often involves removing unusable material in the form of the overburden of soil and rock that covers the desired resource. This may leave a pit that fills with water alongside a pile of the overburden. This kind of mining practice caused many environmental problems in the past. With the modern practice of surface mining, however, topsoil is first removed and stored, rock removed to get to the resource is either placed back in the pit or on contoured piles, and the topsoil placed over it for revegetation. In favorable cases, the result can be attractive lakes that support fish life and vegetated, gently sloping artificial hills.

Underground mining usually does not leave the visible scars that may be inflicted by surface mining. However, it can have profound environmental effects. Collapse of underground mines can cause surface subsidence. Water flowing through and from underground mines can pick up water pollutants. Most ores require a degree of beneficiation in which the usable portion of the ore is concentrated, leaving piles of tailings. These may collapse, and materials leached from them can pollute water. Examples of the latter include acidic water produced by the action of bacteria on iron pyrite, FeS_2, removed from coal and radium leached from the tailings remaining from uranium mining operations.

This page titled 2.5.11.6: Anthrospheric Influences on the Geosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.6: Anthrospheric Influences on the Geosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.11.7: The Geosphere as a Waste Repository

As discussed above, mineral processing produces large quantities of waste solids. Other sources of waste solids include ash from coal combustion, municipal garbage, and solid wastes from various industrial processes. Ultimately, these wastes are placed on or in the geosphere. Such measures have an obvious potential for pollution.

One of the most common waste materials that ends up as part of the geosphere is **municipal refuse**, the "garbage" generated by human activities. This material is largely disposed in **sanitary landfills** made by placing the solid wastes on top of the ground or in depressions in the ground and covering it with soil to minimize effects such as windblown waste paper and plastic, emission of odorous materials to the atmosphere, and water pollution. Although "garbage dumps" used to be notably unsightly and polluting, modern practice of sanitary landfilling can result in areas that can be used as parkland, golf courses, or relatively attractive open space. The unconsolidated nature of decaying garbage and the soil used to cover it make municipal landfills generally unsuitable for building construction. Biological decay of degradable organic material ($\{CH_2O\}$) in the absence of oxygen generates methane gas by a process represented as

$$2\,(CH_2O) \rightarrow CO_2 + CH_4$$

Methane is a powerful greenhouse gas in the atmosphere, much more effective per molecule at absorbing infrared radiation than is CO_2, so it is undesirable to release CH_4 to the atmosphere. However, modern sanitary landfills may be equipped with pipes and collection systems so that the methane can be collected and used as a fuel.

Whereas the release of gases, particularly methane, to the atmosphere is a potential air pollution problem with sanitary landfills, contaminated **leachate** consisting of water seeping through the landfilled wastes can pollute water, especially groundwater. This water may contain heavy metals, organic acids, odor-causing organics, and other undesirable pollutants. There are two general approaches to minimizing problems from contaminated landfill leachate. One of these is to construct the landfill in a manner that minimizes water infiltration, thus reducing the amount of leachate produced to lowest possible levels. To prevent the leachate from getting into groundwater, it is desirable to locate the landfill on a layer of poorly permeable clay. In some cases, the bottom of the landfill may be lined with impermeable synthetic polymer liners that prevent leachate from getting into groundwater. In cases where significant quantities of badly polluted leachate are generated, it is best to collect and treat the leachate, usually by biological treatment processes.

Minimization of the quantities of materials requiring sanitary landfill disposal using the best practice of industrial ecology and green chemistry is highly desirable. The best way to do that is by reducing quantities of materials at the source, simply using less material that ultimately will require disposal. Wherever possible, materials, such as packing materials, that ultimately get into landfills should be biodegradable. Recycling programs in which glass, plastic, paper, and food cans are removed from refuse prior to disposal are effective in reducing quantities of material requiring disposal in landfills. Burning of garbage with proper pollution control measures can reduce it to a low-volume ash that can be placed in a landfill. Although not practiced to a significant extent, anaerobic digestion of macerated wet refuse in an oxygen-free digester has the potential to produce methane for fuel use and greatly reduce the mass of the degradable wastes.

Sanitary landfills are not suitable for the disposal of hazardous substances. These materials must be placed in special **secure landfills**, which are designed to contain the wastes and leachate, thus preventing pollution of water, air, and the geosphere. One way in which this is accomplished is with impermeable synthetic membranes that prevent water from seeping into the fill and prevent leachate from draining into groundwater. These landfills are often equipped with water treatment systems to treat leachate before it is released from the system. Unfortunately, many hazardous chemicals **never** degrade and a "secure" chemical landfill leaves problems for future generations to handle. One of the major objectives of green chemistry is to prevent the generation of any hazardous materials that would require disposal on land. The best way to do that is to avoid making or using such materials. In cases where that is not possible and hazardous materials are generated, they should be treated in a way that renders them nonhazardous prior to disposal.

This page titled 2.5.11.7: The Geosphere as a Waste Repository is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.7: The Geosphere as a Waste Repository** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.11.8: Have You Thanked a Clod Today?

A common bumper sticker is one that asks the question, “Have you thanked a green plant today?” this is an obvious reference to plants whose photosynthesis produces the food that we and most other animals depend upon for our existence. An even more fundamental question is whether we have thanked the soil — the clods of dirt — upon which green plants depend for their existence. Good, productive soil combined with a suitable climate and adequate water is the most valuable asset that a nation can have. Vast areas of the world lack this fundamental asset, and the people living in areas with poor soil often suffer poverty and malnutrition as a result. Furthermore, areas that once had adequate soil have seen it abused and degraded to the extent that it is no longer productive. One of the central challenges faced by the practice of green chemistry and industrial ecology is to retain and enhance the productive qualities of soil.

The remainder of this chapter addresses soil and those aspects of agriculture related specifically to soil. The biological aspects of agriculture and the production of **food and biomass are discussed in Chapter 12.**

What is Soil?

Soil is a term that actually describes a wide range of finely divided mineral matter containing various levels of organic matter and water that can sustain and nourish the root systems of plants growing on it. Soil is largely a product of the weathering of rock by physical, chemical, and biochemical processes that produces a medium amenable to the support of plant growth. A healthy soil contains water available to plants, has a somewhat loose structure with air spaces, and supports an active population of soil-dwelling organisms, including fungi and bacteria that degrade dead plant biomass and animals, such as earthworms. Although the solids in a typical soil are composed of about 95% inorganic matter, some soils contain up to 95% organic matter, and some sandy soils may have only about 1% organic matter.

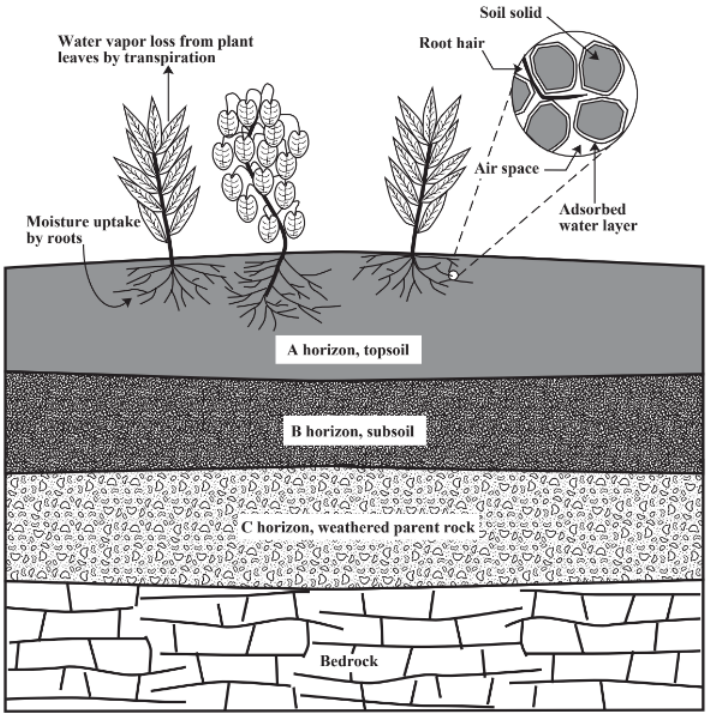


Figure 11.4. Major aspects of soil structure showing various soil horizons. The inset shows aspects of soil microstructure

Figure 11.4 shows the major aspects of the physical structure of soil. Soil is divided into layers called **horizons** formed by weathering of parent rock, chemical processes, biological processes, and the action of water including leaching of colloidal matter to lower horizons. The most important of these for plant growth is **topsoil**. Plant roots permeate the topsoil and take water and plant nutrients from it. Topsoil is the layer of maximum biological activity. The **rhizosphere** is the part of topsoil in which plant roots are especially active and in which the elevated levels of biomass are composed of plant roots and microorganisms associated with them. There are strong synergistic relationships between plant root systems and microorganisms in the rhizosphere. The surfaces of

root hairs are commonly colonized by microorganisms, which thrive upon carbohydrates, amino acids, and root-growth-lubricant mucigel secreted from the roots.

Inorganic Solids in Soil

Reflecting the fact that the two most common elements in the earth's crust are oxygen and silicon (see Section 11.2), silicates are the most common mineral constituents of soil. These include finely divided quartz (SiO_2), orthoclase ($KAlSi_3O_8$), and albite ($NaAlSi_3O_8$). Other elements that are relatively abundant in Earth's crust are aluminum, iron, calcium, sodium, potassium, and magnesium; their abundance is reflected by various minerals such as epidote ($4CaO•3(AlFe)2O_3•6SiO_2•H_2O$), geothite ($FeO(OH)$), magnetite ($Fe_3O_4$), calcium and magnesium carbonates ($CaCO_3$, $CaCO_3•MgCO_3$), and oxides of manganese and titanium in soil. Soil parent rocks undergo weathering processes to produce finely divided colloidal particles, by far the most abundant of which are clays. These secondary minerals hold moisture and mineral nutrients, such as K^+ required for plant growth, that are accessible by plant roots and are repositories of plant nutrients. Inorganic soil colloids can absorb toxic substances in soil, thus reducing the toxicity of substances that would harm plants. It is obvious that the abundance and nature of inorganic colloidal material in soil are important factors in determining soil productivity.

Soil Organic Matter

The few percent of soil mass consisting of organic matter has a strong influence upon the physical, chemical, and biological characteristics of soil. Among its important effects in soil organic matter is effective in holding soil moisture and it holds and exchanges with plant roots some of the ions that are required as plant nutrients. Temperature, moisture, and climatic conditions significantly affect the kinds and levels of soil organic matter. Cold, wet conditions in which soil stays saturated with moisture preventing access of microorganisms to oxygen tend to prevent complete biodegradation of plant residues that compose soil organic matter allowing it to accumulate. This is clearly illustrated by accumulation of peat in Ireland and other locales with similar climatic conditions such that most of the solid soil is composed of organic matter. Tropical conditions, especially with alternate wet and dry seasons, can result in loss of soil organic matter. One reason that the soil supporting tropical rain forests degrades so quickly when the trees are removed is that the organic matter in the soil undergoes rapid biodegradation when the forest cover is removed,

The plant biomass residues that form soil organic matter undergo a biodegradation process by the action of soil bacteria and fungi in which the cellulose in the biomass is readily degraded leaving modified residues of the lignin material that binds the cellulose to the plant matter. This is the process of **humification** and the residue is **soil humus**, a black organic material of highly varied chemical structure. A fraction of soil humus is soluble in water (see the discussion of humic substances in water, Chapter 9, Section 9.3), especially when base is present in the water. Another fraction called **humin** does not dissolve and stays in the solid soil.

Though composing usually not more than a few percent of soil, soil humus has a very strong influence on the characteristics of soil. It has a strong affinity for water and holds much of the water in a typical soil. Primarily because of their carboxylic acid ($-CO_2H$) groups, soil humic molecules exchange H^+ ion and act to buffer the pH of water in soil (the soil solution). Humic substances bind metal ions and other ionic plant nutrients. Soil humus also binds and immobilizes organic materials, such as herbicides applied to soil.

Water in Soil and The Soil Solution

Water in soil is required for plants. This water is taken up by plant root hairs, transferred through the plant, and evaporated from the leaves, a process called **transpiration**. The quantities of water involved are enormous; for example, the water transpired to produce a kilogram of dry hay can amount to several hundred kg. Most of the water in normal soils is not present as visible liquid, but is absorbed to various degrees upon the soil solids. In fact, a condition in which all the spaces in soil are filled with water — waterlogging — slows the growth of most plants. The water that is available in soil is called the **soilsolution** and contains a number of dissolved materials, including plant nutrients. It plays an essential role in transferring substances, such as dissolved metal cations, between roots and the soil solid. Cations commonly present in the soil solution include H^+, Ca^{2+}, Mg^{2+}, Na^+, and K^+ along with very low levels of Fe^{2+}, Mn^{2+}, and Al^{3+}. Commonanions present are HCO_3^-, CO_3^{2-}, HSO_4^-, SO_4^{2-}, Cl^-, and F^-.

This page titled 2.5.11.8: Have You Thanked a Clod Today? is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.8: Have You Thanked a Clod Today?** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.11.9: Plant Nutrients and Fertilizers in Soil

Plant biomass is composed largely of carbon, hydrogen, and oxygen, which plants extract from water and atmospheric carbon dioxide. Other nutrients that plants require in relatively large quantities are calcium, magnesium, and sulfur, which are usually in sufficient abundance in soil, and nitrogen, phosphorus, and potassium, which are commonly added to soil as fertilizers.

Soil acidity in the form of H+ ion builds up as plant roots exchange H+ for other cationic nutrients in soil. When acidity reaches excessive levels, the soil is no longer productive. Acidity can be neutralized by the addition of lime ($CaCO_3$), which neutralizes acidity according to the following reaction:

$$\text{Soil } (H^+)_2 + CaCO_3 \rightarrow \text{Soil) } Ca^{2+} + CO_2 + H_2O$$

This process also adds calcium to soil.

Essential plant nutrient nitrogen is very much involved with nature's nitrogen cycle, which is significantly modified by human activities. Major aspects of this cycle are the following:

- At 79% N_2, Earth's atmosphere constitutes an inexhaustible nitrogen resource, although, because of the extreme stability of the N_2 molecule, it is difficult to extract nitrogen from air in a chemically combined form.
- *Rhizobium* bacteria growing on the roots of leguminous plants, such as clover and soybeans, convert atmospheric nitrogen to nitrogen chemically bound in biomolecules. This nitrogen is converted to ammonium ion, NH_4^+, when plant residues and animal feces, urine, and carcasses undergo microbial decay.
- Lightning and combustion processes convert atmospheric nitrogen to nitrogen oxides, and ammonia manufacturing plants produce $NH_3$from atmospheric elemental nitrogen and elemental hydrogen produced by natural gas.
- Soil microbial processes oxidize ammoniacal nitrogen (NH_4^+) to nitrate ion, NO_3^-, the form of nitrogen most readily used by plants. Microbial processes also produce gaseous N_2 and NO_2 which are released to the atmosphere, a process called denitrification that completes the nitrogen cycle.

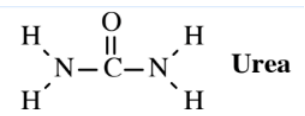
Natural processes usually do not produce sufficient nitrogen to allow maximum plant growth, so that artificial means are used to extract nitrogen in a chemically combined form from the atmosphere. This is done by the Haber process combining elemental N_2 and H_2 over a catalyst at very high pressures of about 1000 times atmospheric pressure and an elevated temperature of 500°C. The reaction is

$$N_2 + 3\,H_2 \rightarrow 2\,NH_3$$

producing ammonia that is 82% chemically bound N. This anhydrous ammonia can be applied directly below the soil surface where its tremendous attraction to soil moisture binds it to the soil. It can also be applied as a 30% solution of NH_3 in water, and is sometimes added directly to irrigation water. Ammonia, which is held in soil as ammonium ion, NH_4^+, is not well assimilated directly by most plants. But it is slowly oxidized by the action of soil bacteria using atmospheric O_2 oxidant to nitrate ion, NO_3^-, which is used directly by plants.

A solid form of nitrogen fertilizer can be made by reacting ammonia with oxygen over a platinum catalyst to make nitric acid, HNO_3, and reacting the acid with basic ammonia to make ammonium nitrate, NH_4NO_3. This molten material is solidified into small pellets that can be applied to soil as fertilizer. Ammonium nitrate mixed with fuel oil is used for blasting to quarry rock, and it was the explosive used in the bombing of the Oklahoma City Federal Building in 1995. A safer alternative to ammonium nitrate as a solid nitrogen fertilizer is urea, which is made by a process that, overall, involves the reaction of carbon dioxide and ammonia:

$$CO_2 + 2\,NH_3 \rightarrow CO(NH_2)_2 + H_2O$$



Urea

Phosphorus is an essential plant nutrient required for cellular DNA and other biomolecules. Itis utilized by plants as $H_2PO_4^-$ and HPO_4^{2-} ions. Phosphate minerals that can be used to manufacture phosphorus-containing fertilizers occur in a number of places throughout the world. In the United States, Florida has especially abundant phosphate resources, largely as fluorapatite,

$Ca_5(PO_4)_3F$, as well as hydroxyapatite, $Ca_5(PO_4)_3OH$. These phosphate minerals are too insoluble to serve directly as fertilizers and are treated with phosphoric acid and sulfuric acid to make superphosphates that are much more soluble and available to plants:

$$2\,Ca_5(PO_4)_3F\,(s) + 14\,H_3PO_4 + 10\,H_2O \rightarrow 2\,HF\,(g) + 10\,Ca(H_2PO_4)_2 \cdot H_2O$$

$$2\,Ca_5(PO_4)_3F\,(s) + 7\,H_2SO_4 + 3\,H_2O \rightarrow 2\,HF\,(g) + 3\,Ca(H_2PO_4)_2 \cdot H_2O + 7\,CaSO_4$$

Potassium as the potassium ion, K^+, is required by plants to regulate water balance, activate some enzymes, and enable some transformations of carbohydrates. Potassium is one of the most abundant elements in the earth's crust, of which it makes up 2.6%; however, much of this potassium is not easily available to plants. For example, some silicate minerals such as leucite, $K_2O{\bullet}Al_2O_3{\bullet}4SiO_2$, contain strongly bound potassium. Exchangeable potassium held by clay minerals is relatively more available to plants. Potassium for fertilizer is simply mined from the ground as salts, particularly KCl, or pumped from beneath the ground as potassium-rich brines. Large potassium deposits occur in the Canadian province of Saskatchewan.

Plants require several **micronutrients**, largely elements that occur only at trace levels, for their growth. These include boron, chlorine, copper, iron, manganese, molybdenum (for N-fixation), and zinc. Some of these are toxic at levels above those required for optimum plant growth. Most of the micronutrients are required for adequate function of essential enzymes. Photosynthetic processes use manganese, iron, chlorine, and zinc. Since the micronutrients are required at such low levels, soil normally provides sufficient amounts.

This page titled 2.5.11.9: Plant Nutrients and Fertilizers in Soil is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.9: Plant Nutrients and Fertilizers in Soil** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.11.10: Soil and Plants Related To Wastes and Pollutants

Soil is a repository of large quantities of wastes and pollutants, and plants act as filters to remove significant quantities of pollutants from the atmosphere. Sulfates and nitrates from the atmosphere, including acid-rain-causing H_2SO_4 and HNO_3 deposit largely on the land and the plants growing on it. Gaseous atmospheric SO_2, NO, and NO_2 are absorbed by soil and oxidized to sulfates and nitrates. Soil bacteria and fungi are known to convert atmospheric CO to CO_2. When leaded gasoline was widely used, soil along highways became contaminated with lead, and lead mines and smelters were significant sources of this toxic element. Organic materials, such as those involved in photochemical smog formation, are removed by contact with plants and are especially attracted by the waxy organic-like surfaces of the needles of pine trees.

A number of materials that can be considered as pollutants are deliberately added to soil. The most obvious of these consists of insecticides and herbicides added to soil for pest and weed control. Chemicals from hazardous waste disposal sites can get onto soil or below the soil surface by leaching from landfill or drainage from waste lagoons. Some kinds of wastes, especially petroleum hydrocarbons, are disposed on soil where adsorption and microbial processes immobilize and degrade the wastes. Soil can be effective for the treatment of sewage. Leakage from underground storage tanks of organic liquids, such as gasoline and diesel fuel, have created major soil contamination problems.

Soils in parts of New York State have been contaminated with polychlorinated biphenyls (PCBs) discarded from the manufacture of industrial capacitors. Analyses of PCBs in United Kingdom soils archived for several decades have shown levels of these pollutants that parallel their production. Starting with very low levels around 1940 before PCBs were manufactured in large quantities, concentrations of PCBs increased markedly, peaking around 1970, when PCB manufacture was ceased. More recent soil samples have shown PCB concentrations near the pre-1940 levels. It is believed that these results reflect evaporation of PCBs and their condensation onto soil. They are consistent with observations of high PCB levels in remote Arctic and sub-Arctic regions believed to be due to the condensation of these compounds from the atmosphere onto soil in very cold regions.

The degradation and eventual fates of the enormous quantities of herbicides and other pesticides applied to soil are very important in understanding the environmental effects of these substances. Many factors are involved in determining pesticide fate. One of the main ones of these is the degree of adsorption of pesticides to soil, strongly influenced by the nature and organic content of the soil surface as well as the solubility, volatility, charge, polarity, and molecular structure and size of the pesticides. Strongly adsorbed molecules are less likely to be released and thus harm organisms, but they are less biodegradable in the adsorbed form. The leaching of adsorbed pesticides into water is important in determining their water pollution potential. The effects and potential toxicities of pesticides to soil bacteria, fungi, and other organisms have to be considered. It must be kept in mind that pesticides may be converted to more toxic products by microbial action.

This page titled 2.5.11.10: Soil and Plants Related To Wastes and Pollutants is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.10: Soil and Plants Related To Wastes and Pollutants** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.11.11: Soil Loss - Desertification and Deforestation

Soil erosion refers to the loss and relocation of topsoil by water and wind action. About a third of U.S. topsoil has been lost to erosion since cultivation began on the continent and at present about a third of U.S. cropland is eroding at a rate sufficient to lower productivity. About 10% of U.S. land is eroding at an unacceptable rate in excess of 14 tons of toposil per acre annually. Soil erosion is largely a product of cultivation. Except in cases of extreme slopes, very high winds, and torrential rains, uncultivated soils undergo little erosion. Erosion was recognized as a problem in the central United States within a few years after forests and prairie grasslands were first plowed to plant crops, particularly in the latter 1800s. The recognition that precious topsoil was being lost at an unsustainable rate lead to soil conservation measures going back to 1900, or even earlier. In that sense, soil conservation was the first environmental movement, predating efforts to alleviate water and air pollution by many decades.

Water erosion is responsible for greater loss of soil than is wind erosion. Whereas wind erosion tends to move soil around and deposit it in areas where it can still be used for growing crops, water erosion normally moves greater quantities of soil and carries them into streams and rivers and ultimately to the oceans. The overall pattern of soil erosion in the Central Continental United States is shown in Figure 11.5. This figure shows that erosion is especially bad in agricultural areas draining into the Missouri and Mississippi Rivers; millions of tons of soil are carried by these rivers into the Gulf of Mexico each year. These are areas of relatively high rainfall, which can sometimes come as very intense rainstorms, especially during the spring. A high proportion of the farmlands in these areas are devoted to row crops, which are crops such as corn, soybeans, and sorghum grains planted in rows with bare soil in between. This mode of cultivation leaves soil that is especially susceptible to water erosion.

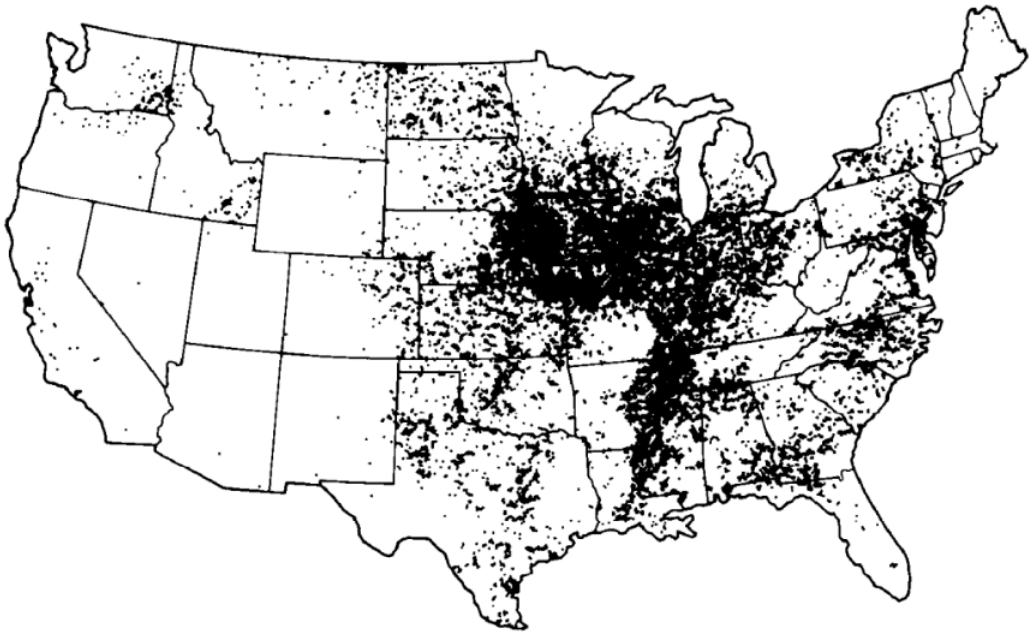


Figure 11.5 Pattern of soil erosion (dark areas) in the major agricultural states of the central continental U.S. It is seen that erosion is particularly severe in the highly cultivated Missouri and Mississippi River watersheds.

The ultimate result of soil erosion and other unsustainable agricultural practices in relatively dry areas is **desertification**. This condition occurs when permanent plant cover is lost from soil so that it loses its capacity to retain moisture, dries out, and loses fertility so that plants no longer grow on it. Among the interrelated factors involved in desertification are wind erosion, water erosion (which occurs during sporadic cloudbursts even in arid areas), development of adverse climate conditions, depletion of underground water aquifers, lack of water for irrigation, accumulation of salt in water supplies, loss of soil organic matter, and deterioration of soil physical and chemical properties. Eventually the land becomes unable to support agriculture, grazing, or even significant human populations. Desertification is one of the most troublesome results of global warming caused by greenhouse gases. It is actually a very old problem and is a serious concern in many parts of the world, such as the Mideast, the southern boundary of Africa'sSahara, and regions of the southwestern U. S. Formerly productive areas of the Middle East and North Africa, "lands of milk and honey" described in biblical terms, have turned into desert, largely due to human agricultural activities. The growth of domestic grazing animals on these areas — especially goats, which tend to pull vegetation up by its roots — has been a

particularly strong contributor to desertification. Much of the productive capacity of arid grasslands in the western and southwestern United States has been drastically diminished by overgrazing.

Fortunately, human ingenuity and technological tools can be used to prevent or reverse desertification. For example, water, which upon occasion falls as torrential rain upon normally dry desert lands, can be collected and used to recharge underground water aquifers. Advanced cultivation and irrigation techniques can be used to establish perennial plant cover on erosion-prone desert soils. Potentially, plants can be genetically engineered to grow under severe conditions of temperature, drought, and salinity. Environmentally friendly mining practices can be employed to obtain minerals, and land surfaces damaged by harmful strip mining practices can be restored.

The loss of forest growth to cultivated land —**deforestation**— has occurred extensively in the United States. However, much of the colonial U.S., particularly in New England, which was deforested for cultivation of crops, is now undergoing largely spontaneous reforestation as unprofitable farmlands are abandoned and trees become established again. Deforestation is a particularly severe problem in tropical regions. Rich tropical forests contain most known plant and animal species many of which are becoming extinct as the forests are destroyed. Once destroyed, tropical forests are almost impossible to restore. This is because tropical forest soil has been leached of nutrients by the high annual rainfalls in tropical regions. When forest cover is removed, the soil erodes rapidly, loses the plant roots and other biomass that tends to hold it together, loses nutrients, and becomes unable to sustain either useful crops or the kinds of forests formerly supported.

The key to preventing soil loss from erosion as well as preventing desertification from taking place lies in a group of practices that agriculturists term **soil conservation**. A number of different approaches are used to retain soil and enhance its quality. Some of these are old, long-established techniques such as construction of terraces and planting crops on the contour of the land (see Figure 11.6). Crop rotation and occasional planting of fields to cover crops, such as clover, are also old practices. A relatively new practice involves minimum cultivation and planting crops through the residue of crops from the previous year. This practice, now commonly called **conservation tillage**, is very effective in reducing erosion because of the soil cover of previous crops and the roots that are left in place. Conservation tillage does make use of herbicides to kill

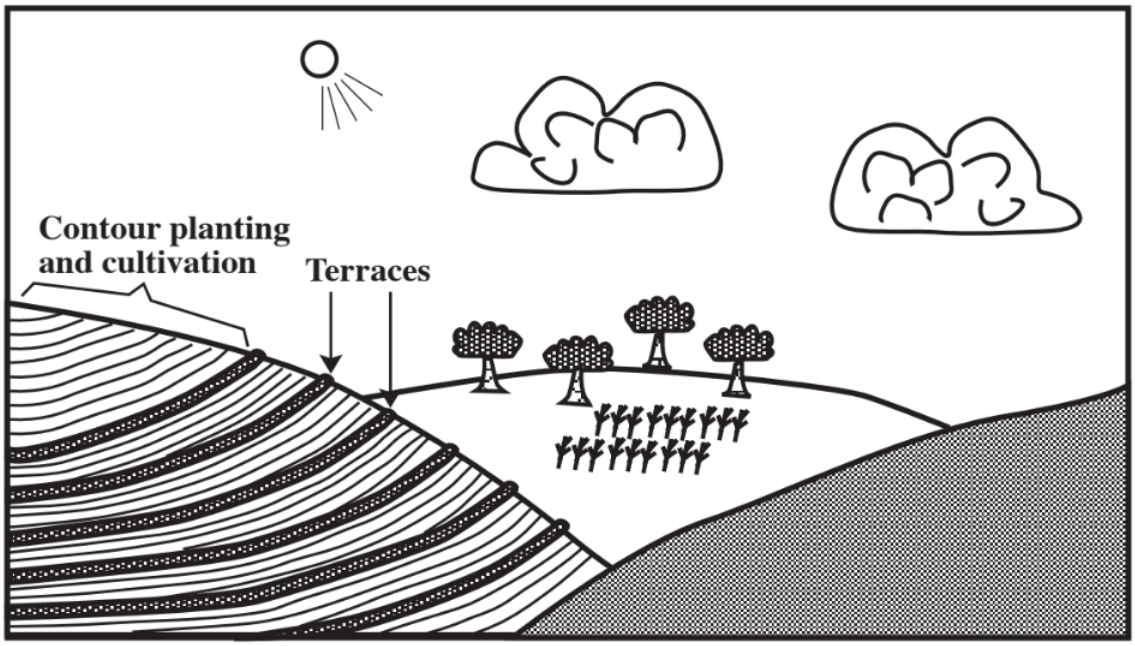


Figure 11.6.
Construction of terraces on the contour of land and planting crops on the contour are practices that have been very effective in reducing soil erosion

competing weeds until the desired crop is established enough to shade out competing plants, but only minimum quantities of herbicides are applied. There is some concern that fungi (molds) will thrive in old crop residues and cause problems with new crops.

The ultimate in no-till agriculture is the use of perennial plants that do not have to be planted each year. Trees in orchards and grape vines in vinyards are such plants. The roots of perennial plants are very effective in holding soil in place. Efforts to develop perennial plants that produce grain have not been successful to date. This is because a successful grain-producing plant is one that dedicates its metabolic processes to the production of large quantities of seed that can be used for grain, whereas perennial plants put their energy into the development of large, bulbous root structures that store food for the next growing system. It is possible that sometime in the future genetic engineering may be applied to the development of perennial plants capable of producing high grain yields.

Among the most successful plants at stopping erosion are trees. These plants grow for many years and some tree varieties will grow back from their root structures when the wood is harvested. Wood and wood products are probably the most widely used renewable resources. Hybrid tree varieties have been developed that are outstanding producers of biomass.

Wood is a renewable resource used for many purposes. In construction, wood substitutes for steel, aluminum, and cement. All of these materials are produced by very energy-intensive processes, so substitution of wood, where applicable, conserves large amounts of energy. Wood is about 50% cellulose, a carbohydrate polymer that is used directly to make paper. Although humans and many other animals cannot use cellulose directly for fuel, it can be broken down chemically or biochemically to glucose sugar. This material serves as a food source for yeasts (a form of fungi) that generate ethanol, an alcohol that can be used as fuel and to make other chemicals. In the process, the yeasts produce protein that can be fed to animals.

Biochar for Soil Conservation and Enrichment

A relatively recent development in enhancing soil quality is the use of biochar as a soil amendment. **Biochar** is made by the pyrolysis of organic matter such as crop residues. It is a natural constituent of soil as the result of forest and prairie fires, but is now produced artificially for addition to soil.[1] The two major advantages of biochar are its high affinity for nutrients by adsorption and its extremely high persistence; unlike humic material it never degrades. Furthermore, the production of biochar has the net effect of permanently sequestering atmospheric carbon dioxide fixed by photosynthesis thus helping to alleviate global warming.

Water and Soil Conservation

Conservation of soil and conservation of water go together very closely. Soil is normally the first part of the geosphere that water contacts, and contaminated soil yields contaminated water. Most fresh water falls initially on soil, and the condition of the soil largely determines the fate of the water and how much is retained in a usable condition. Soil in a condition that retains water allows rainwater to infiltrate into groundwater. If water drains too rapidly from soil, the soil erodes and the water runoff is badly contaminated with soil sediments. Measures taken to conserve soil usually conserve water as well. Terraces, contour cultivation, constructed waterways, and water-retaining ponds (Figure 11.6) prevent water from washing soil away, but also retain water and help prevent flash floods. Some of these measures, especially terracing, involve modification of the contour of the soil. Bands of trees can be planted on the contour to retain both soil and water. Reforestation of land unsuitable for growing crops and avoiding practices, such as overgrazing, that tend to lead to desertification conserve water as well as land.

This page titled 2.5.11.11: Soil Loss - Desertification and Deforestation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **11.11: Soil Loss - Desertification and Deforestation** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.11.12: Literature Cited and Supplementary References

LITERATURE CITED

1. Lehmann, Johannes, and Stephen Joseph, *Biochar for Environmental Management: Science and Technology*, Earthscan, London, 2009.

SUPPLEMENTARY REFERENCES

Baker C. J., and K. E. Saxton, Eds., *No-Tillage Seeding in Conservation Agriculture*, 2nd ed., Cabi Publishing, Wallingford, UK, 2007.

Boardman, John, and Jean Poesen, Eds., *Soil Erosion in Europe*, Wiley, Hoboken, NJ, 2006.

Brady, Nyle C., and Ray R. Weil, *Elements of The Nature and Properties of Soils*, 3rd ed., Pearson Education, Upper Saddle River, NJ, 2010.

Bryant, Edward, *Natural Hazards*, 2nd ed., Cambridge University Press, Cambridge, U.K., 2005.

Chamley, Herve, *Geosciences, Environment and Man*, Elsevier, New York, 2003.

Chesworth, Ward, Ed., *Encyclopedia of Soil Science*, Springer, Dordrecht, Netherlands, 2008.

Coch, Nicholas K., *Geohazards: Natural and Human*, Prentice Hall, Upper Saddle River, NJ, 1995.

Coyne, Mark S., and James A. Thompson, *Fundamental Soil Science*, Thomson Delmar Learning, Clifton Park, NY, 2006.

De Boer, Jelle Zeilinga, and Donald T. Sanders, *Volcanoes in Human History: The Far Reaching Effects of Major Eruptions*, Princeton University Press, Princeton, NJ, 2001.

Eash, Neal S., *Soil Science Simplified*, 5th ed., Blackwell Publishing, Ames, IA. 2008.

Eby, G. Nelson, *Principles of Environmental Geochemistry*, Thomson-Brooks/Cole, Pacific Grove, CA, 2004.

Essington, Michael E., *Soil and Water Chemistry: An Integrative Approach*, Taylor & Francis/CRC Press, Boca Raton, FL, 2004.

Ferry, Natalie, and M. R. Angharad M.R., Eds., *Environmental Impact of Genetically Modified Crops*, Gatehouse, Cambridge, MA, 2009.

Gardiner, Duane T., and Raymond W. Miller, *Soils in Our Environment*, 11th ed., Pearson/Prentice Hall, Upper Saddle River, NJ, 2008.

Gates, Alexander E., and David Ritchie, *Encyclopedia of Earthquakes and Volcanoes*, 3rd ed., Facts on File, New York, 2007.

Hyndman, Donald, and David Hyndman, *Natural Hazards and Disasters*, 2nd ed., Thomson Brooks/Cole, Belmont, CA, 2009.

Lal, Rattan, S*oil Degradation in the United States: Extent, Severity, and Trends*, CRC Press/Lewis Publishers, Boca Raton, FL, 2003.

Magdoff, Fred, and Harold van Es, *Building Soils for Better Crops: Sustainable Soil Management*, Sustainable Agriculture Research and Education Program, Beltsville, MD, 2009.

McCollum, Sean, *Volcanic Eruptions, Earthquakes, and Tsunamis*, Chelsea House, New York,2007.

Marti, Joan, and Gerald Ernst, Eds., *Volcanoes and the Environment*, Cambridge University Press, Cambridge, U.K., 2005.

Montgomery, Carla W., *Environmental Geology*, 8th ed., McGraw-Hill, Boston, 2008.

Morgan, R. P. C., *Soil Erosion and Conservation*, 3rd ed., Blackwell Publishing, Malden, MA, 2005.

Nash, David J., and Sue J. McLaren, Eds., *Geochemical Sediments and Landscapes*, Blackwell Publishing, Malden, MA, 2007.

Paul, Eldor A., *Soil Microbiology, Ecology, and Biochemistry*, 3rd ed., Academic Press, Boston,2007.

Pipkin, Bernard W., *Geology and the Environment*, 5th ed., Thomson Brooks/Cole, Belmont, CA,2008.

Plaster, Edward J.,*Soil Science and Management*, 5th ed., Delmar Cengage Learning, Clifton Park, NJ, 2009.

Sammonds, P. R., and J. M. T. Thompson, Eds., *Advances in Earth Science: From Earthquakes to Global Warming*, Imperial College Press, London, 2007.

Savino, John, and Marie D. Jones, *Supervolcano: The Catastrophic Event that Changed the Course of Human History (Could Yellowstone be Next?)*, New Page Books, Franklin Lakes, NJ, 2007.

Singer, Michael J. and Donald N. Munns, *Soils: An Introduction*, 6th ed., Pearson Prentice Hall, Upper Saddle River, NJ, 2006.

Sparks, Donald L., *Environmental Soil Chemistry*, 2nd ed., Academic Press, Boston, 2003.

Stone, George W.,*Raging Forces: Life on a Violent Planet*, National Geographic, Washington, DC, 2007.

Tabatabai, M. A., and D. L. Sparks, Eds., *Chemical Processes in Soils*, Soil Science Society of America, Madison, WI, 2005.

Van Elsas, Jan Dirk, Janet K. Jansson, and Jack T. Trevors, Eds., *Modern Soil Microbiology*, 2nded., Taylor & Francis/CRC Press, Boca Raton, FL, 2007.

Van der Pluijm, Ben A., and Stephen Marshak, *Earth Structure: An Introduction to Structural Geology and Tectonics*, W.W. Norton, New York, 2004.

Waltham, Tony, *Foundations of Engineering Geology*, Taylor and Francis, London, 2009.

White, Robert E., *Principles and Practice of Soil Science: The Soil as a Natural Resource*, 4thed., Blackwell Publishing., Malden, MA, 2006.

2.5.11.12: Literature Cited and Supplementary References is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Literature Cited and Supplementary References** has no license indicated.

2.5.11.13: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. Suggest the main contributions made by the geosphere to the biosphere.

2. Distinguish between rocks and minerals.

3. How does igneous rock turn into secondary minerals?

4. What is the branch of chemistry that deals with rocks and minerals and their chemical characteristics and interactions?

5. Suggest ways in which improved materials, some made by green chemical processes, can reduce the effects of earthquakes.

6. In what respect do volcanoes have the potential to drastically affect global climate? Is there any evidence for such an effect?

7. How may human activities lead to landslides?

8. What are the formations called that contain water under ground? What is a major threat to groundwater in such formations?

9. What is FeS_2? Why is its exposure of this material from mining a potential problem?

10. Give an example of an air pollutant and of a water pollutant that may be generated by sanitary landfills.

11.How is soil divided? Which is the top one of these divisions?

12. What is humification, and what does it have to do with soil?

13. What is water in soil called? Give the name of the process by which this water enters the atmosphere by way of plants.

14. In what respects is conservation tillage consistent with the practice of green chemistry?

15. Name a gaseous form and two solid forms of fixed nitrogen used as fertilizer.

16. Explain what is meant by desertification.

17. What is the good news in the U. S. regarding deforestation?

18. What is the potential use of perennial plants in grain production and how does it tie in with soil conservation?

19. What are the processes occurring in soil that operate to reduce the harmful effects of pollutants?

20.Of the following, the one that is **not** a manifestation of desertification is (explain): (A) Declining groundwater tables, (B) salinization of topsoil and water, (C) increased organic matter in soil, (D) reduction of surface waters, (E) unnaturally high soil erosion.

21.Why do silicates and oxides predominate among earth's minerals?

22. Explain how the following are related: weathering, igneous rock, sedimentary rock, soil.

23.What is the distinction between weathering and erosion? Suggest ways in which air pollution may contribute to both phenomena.

23.In what respect is biochar a material that gets into soil by natural processes? Where would soils with significant levels of naturally-occurring biochar likely be found?

24.In what sense may volcanoes contribute to air pollution? What possible effects may this have on climate?

25.Large areas of central Kansas have vast deposits of halite. What is halite? What does this observation say about the geologic history of the area?

26. One way in which coal and other fossil fuels may be used without contributing to higher levels of greenhouse gas carbon dioxide in the atmosphere is through carbon sequestration by pumping carbon dioxide into mineral strata. Explain with a chemical

reaction how formations of limestone (calcium carbonate) might be used for this purpose. Suggest how this might cause problems on the surface.

27. In what respects do humus and biochar perform similar functions in soil? What are the main differences between these two kinds of materials.

28. Explain how bamboo might be used for restoring degraded soil. How does bamboo prevent erosion?

2.5.11.13: Questions and Problems is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** has no license indicated.

SECTION OVERVIEW

2.5.12: The Biosphere and the Role of Green Chemistry in Feeding a Hungry World

“From the 1990s, genetically engineered crops resistant to glyphosate herbicide that can be sprayed directly for weed control resulted in a revolution in the production of corn, soybeans, and cotton. Unfortunately, weeds are now emerging that are resistant to glyphosate. Pigweed has been an especially bad actor that can grow 7 or 8 centimeters in a day, reach heights of 2 meters, and with stalks so thick and strong that they can damage harvesting machinery.”

2.5.12.1: Pigweed’s Revenge

2.5.12.2: The Biosphere

2.5.12.3: Cells - Basic Units of Life

2.5.12.4: Metabolism and Control in Organisms

2.5.12.5: Reproduction and Inherited Traits

2.5.12.6: Stability and Equilibrium of the Biosphere

2.5.12.7: DNA and the Human Genome

2.5.12.8: Genetic Engineering

2.5.12.9: Biological Interaction with Environmental Chemicals

2.5.12.10: Biodegradation

2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture

2.5.12.12: Agricultural Applications of Genetically Modified Organisms

2.5.12.13: The Anthrosphere in Support of the Biosphere

2.5.12.14: Livestock and their Wastes

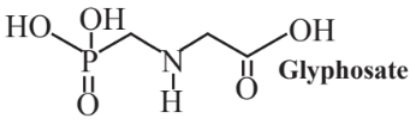
2.5.12.15: Questions and Problems

2.5.12.16: Supplementary References

This page titled 2.5.12: The Biosphere and the Role of Green Chemistry in Feeding a Hungry World is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.12.1: Pigweed's Revenge

Glyphosate, marketed by Monsanto under the brand name Roundup, has been an ideal herbicide that is effective against a broad spectrum of weeds, with low toxicity to animals, and readily degraded in the environment. In the 1990s Monsanto began selling "Roundup ready" seeds of corn, soybeans, and cotton genetically engineered to resist the herbicidal action of glyphosate. These crops could be sprayed directly with glyphosate, killing competing weeds and leaving the crops untouched. This enabled a revolution in agriculture eliminating the need to till the crops and facilitating the adoption of environmentally friendly conservation (no till) crop production, saving large amounts of fuel formerly consumed in tillage. By 2010 in the U.S. about 90% of the soybeans and 70% of corn and cotton were glyphosate-resistant varieties.



Glyphosate

Unfortunately, not long after the introduction of Roundup ready seeds, glyphosate-resistant weeds began appearing including horseweed, giant ragweed, and several of a number of species in the genus *Amaranthus* (pigweed) that by 2010 had afflicted 7-10 million acres of the approximately 170 million acres of corn, soybeans, and cotton planted in the U.S.. Pigweed has been an especially bad actor that can grow 7 or 8 centimeters in a day, reach heights of 2 meters, and with stalks so thick and strong that they can damage harvesting machinery. The first glyphosate-resistant weeds appeared in California in year 2000. In 2003 in the U.S. such weeds had been observed in 10 states, in 2006 14 states and in 2009 20 states. By 2010 there were known to be 10 glyphosate-resistant weed species infesting 7-8 million acres of soybeans, cotton, and corn in 22 states.

Dealing with glyphosate-resistant weeds is a major challenge to the agricultural industry. In response to this problem crop varieties resistant to other herbicides including glufosonate, Dicamba, Sygenta's Calliston, and 2,4-D are being developed.

It is interesting to note that some of the amaranth species that include pigweed have the potential to serve as a major food source with leaves that can be consumed as leafy green vegetables and seeds that can be made into a protein-rich flour. Now sold in some health food stores, amaranth has been a staple food in some native cultures in Peru and Mexico.

This page titled 2.5.12.1: Pigweed's Revenge is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.1: Pigweed's Revenge** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.2: The Biosphere

The **biosphere** consists of all living organisms and the materials and structures produced by living organisms. There is a very close connection between the biosphere and green chemistry including the following:

- Living organisms produce a wide range of materials that are used by humans for a variety of purposes.
- Large quantities of substances including pesticides and fertilizers are generated in the anthrosphere for use to control pests and enhance the growth and health of organisms in the biosphere.
- Reduction of the use and generation of toxic substances in the anthrosphere is designed to prevent harm to humans and other organisms in the biosphere.
- Environmental conditions largely determined by anthrospheric activities strongly affect organisms in the biosphere

Individual organisms in the biosphere and organisms interacting in ecosystems can teach humans a lot about how to apply green chemistry. One important respect in which this is done is by the mild conditions under which organisms carry out complex chemical syntheses. Living organisms can function only within narrow temperature ranges that are close to those that humans find comfortable. (Even the 90-100° C temperatures under which thermophilic bacteria function in hot springs and similar locations are not very far from room conditions.) Therefore, the enzyme-catalyzed reactions that organisms carry out occur under much milder conditions than the often high-temperature, high-pressure conditions of conventional chemical synthesis. Furthermore, organisms cannot tolerate highly toxic substances that are often used in chemical synthesis, the elimination of which is a primary objective of the practice of green chemistry.

Another lesson that living organisms provide for an efficiently operating anthrosphere is in the relationships between organisms with each other and with their environment in biological ecosystems. The wide variety of such ecosystems that have evolved over hundreds of millions of years of evolution have had to be sustainable to survive, completely recycling materials and preserving and enhancing their environment. This is in contrast to the way in which anthrospheric systems have evolved, especially during the last two centuries of the industrial revolution. In general, humans and their industrial systems have exploited nonrenewable resources and have polluted water, air, and land in a manner that simply cannot be sustained. Humans have a lot to learn from the biosphere regarding how to operate the anthrosphere sustainably (see the discussion of industrial ecology in Chapter 13) in which various enterprises compose sustainable industrial ecosystems analogous to ecosystems in the biosphere.

Biology

Biology is the science of life and the organisms that comprise life. So what is life? Biologists define living organisms as those that share (1) constitution by particular classes of life molecules,(2) hierarchical organization, (3) capability to carry out metabolic processes, (4) ability to reproduce, (5) development, and (6) heredity

The kinds of molecules that comprise living organisms were discussed in Chapter 7. Recall that these are *proteins* composed of polymers of nitrogen-containing amino acids, *carbohydrates* consisting of small molecules and polymers with an approximate simple formula of CH_2O, *lipids* defined by their property of solubility in organic solvents, and *nucleicacids* that are long polymers of sugars, nitrogen-containing bases, and phosphate. Two of these kinds of materials are often bonded together as hybrid molecules. Along with water and some kinds of salts they make up the matter in living organisms. Literally thousands of kinds of structural and functional characteristics are possible with the four kinds of molecules mentioned above. For example, proteins comprise muscle tissue and make up the enzyme molecules that act as catalysts to enable biochemical reactions to occur. A simple carbohydrate, glucose, $C_6H_{12}O_6$, is the primary organic product generated by plant photosynthesis and is present in animal bloodstreams. Large numbers of glucose molecules bonded together make up polymeric cellulose that is the structural material in plants. Lipids make up the crucial membranes that enclose living cells. And nucleic acids compose the genetic material that regulate cell function and reproduction.

Hierarchical organization applies to living organisms from the level of atoms all the way to the biosphere as a whole. Proteins, carbohydrates, lipids, and nucleic acids in living organisms are organized into distinct microscopic bodies contained in cells and called **organelles**. Cells are bodies of several micrometers (μm) in size that are the basic building blocks of organisms in that they are the smallest bodies of organisms that can exist independently (even cells of humans can be grown in cell cultures outside the body, given the appropriate nutrients and conditions). In higher organisms cells with similar functions comprise **tissues** and tissues in turn make up **organs**, which may be organized into whole systems of organs. An **organism** is a collection of organs and organ

systems. Organisms from the same species assembled in a group comprise a **population** and a cluster of populations existing in the same place makes up a **community**. Numerous communities living in a particular environmental area, interacting with each other and with their environment, make up an **ecosystem**. Finally, all Earth's ecosystems comprise the entire **biosphere.**

The process of **metabolism** is what occurs when organisms mediate chemical (biochemical)processes to get energy, make raw materials required for tissues in organisms or modify raw materials for this purpose, and reproduce. Although there are thousands of different metabolic reactions, two stand out. The first of these is *photosynthesis* shown in Reaction 12.3.1 in which plants use light energy to convert inorganic CO_2 and H_2O to glucose sugar, $C_6H_{12}O_6$. The second major type of metabolic reaction is the mirror image of photosynthesis, **cellular respiration** in which glucose is oxidized to CO_2 and H_2O, yielding energy that is used by the organism. An interesting aspect of the conversion and utilization of energy in metabolism is that all organisms use the high-energy chemical species **adenosine triphosphate**, **ATP**, (structural formula shown in Chapter 7, Section 7.8) to transfer, convert, and store energy.

All organisms undergo **reproduction** to produce offspring to continue the species. In addition to continuing a species, reproduction enables evolution to occur that results in new species.

Development is the process that occurs as an organism progresses from a fertilized egg to a juvenile and on to adulthood. Development occurs in higher forms of life (obvious in human babies) and even single-celled bacteria that reproduce by cell division undergo development as the cells grow and produce additional organelles prior to further division.

Heredity refers to the process by which traits characteristic of a species of organism are passed on to later generations. Heredity occurs through the action of DNA. Heredity is the mechanism by which organisms have undergone evolution and adaptation to their environment.

Organisms that comprise living beings in the biosphere range in size and complexity from individual bacterial cells less than a micrometer in dimensions up to giant whales and human beings capable of thought and reasoning and may be divided into several kingdoms. **Archaebacteria** and **Eubacteria** are generally single-celled organisms without distinct, defined nuclei. **Protists** are generally single-celled organisms that have cell nuclei and may exhibit rather intricate structures. The three other kingdoms are **Plantae**(plants), **Animalia** (animals), and **Fungi** typified by molds and mushrooms.

Organisms are classified according to their food and energy sources and their utilization of oxygen. **Autotrophs** synthesize their food and biomass from simple inorganic substances, usually using solar energy to perform **photosynthesis**. **Chemautotrophs** mediate inorganic chemical reactions for their energy. **Heterotrophs**, including humans, derive their energy and biomass from the metabolism of organic matter, usually biomass from plants. **Oxic** (aerobic) organisms require oxygen, whereas **anoxic** (anaerobic) organisms use alternate sources of oxidants. **Facultative** organisms can use oxygen or other oxidants depending upon conditions.

The biosphere is greatly influenced by the other environmental spheres. In an environment where temperatures are moderate, sunshine abundant, and nutrients readily available, the biosphere consists of diverse groups of organisms interacting and codependent within thriving ecosystems. Under extreme conditions, there may be only a few organisms composing the localized biosphere, specialized for existence at extreme temperatures, high acidities, high levels of pollutants, or other conditions that make life impossible for most organisms.

Just as the biosphere is strongly influenced by the environment in which the organisms are found, it has a strong influence upon its surroundings. Organisms act to break down inhospitable rock to form soil that supports a variety of plants. The oxygen in the atmosphere, upon which we all depend for our existence, was put there by photosynthesis performed by bacteria capable of photosynthesis. The nature of the anthrosphere that humans have constructed is influenced by the biosphere; examples include dwellings constructed of wood, and shelters and clothing used by Plains Native Americans made from bison hides.

This page titled 2.5.12.2: The Biosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.2: The Biosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.3: Cells - Basic Units of Life

As a fundamental unit of the biosphere, it is appropriate to choose living **cells**, which were discussed in Chapter 7, Section 7.2, as entities in which biochemical processes occur. A single cell visible only under a microscope may perform all the functions required for an organism to process nutrients and energy and to reproduce. Or cells may be highly specialized entities, such as human liver, brain, and red blood cells.

Cell structure has an important influence on determining the nature of the biomaterials generated by biochemical processes in the cells. Muscle cells consist largely of strong structural proteins capable of contracting and movement. Bone cells secrete a protein mixture that then mineralizes with calcium and phosphate to produce solid bone. The walls of cells in plants are largely composed of strong cellulose, which makes up the sturdy structure of wood.

As noted in Section 7.2, there are two general classes of cells. **Prokaryotic cells** are those that make up bacteria and simple single-celled organisms that composed all of life on Earth for the first approximately 2 billion years of life on the planet. These cells are only about 1–2 micrometers in size, have only limited external appendages, and have relatively less (though still complex)internal structures. **Eukaryotic cells** compose all organisms other than bacteria, are typically 10μm or more in size, often have external appendages, and generally show well differentiated internal structures with numerous distinct parts. These cells appeared only about 1.5 billion years ago in the estimated 3.5 billion years that life has existed on Earth. Figure 12.1 represents prokaryotic cells and plant and animal eukaryotic cells.

Three features largely distinguish plant eukaryotic cells from animal cells in that the plant cells have a cell wall, a large central vacuole, and chloroplasts. The cell wall gives the plant cell strength and rigidity. The vacuole takes up most of the cell volume and allows contact with gases. The chloroplasts are sites in which chlorophyll uses light energy ($h\nu$) to synthesize carbohydrates as shown by the following reaction for the photosynthetic generation of glucose sugar:

$$6\,CO_2 + 6\,H_2O + (\text{light energy}, h\nu) \rightarrow C_6H_{12}O_6\ (\text{glucose}) + 6\,O_2$$

Shown by the above reaction, **photosynthesis** was responsible for the greatest changes that the biosphere has ever caused in the atmosphere and geosphere. This occurred with the evolution of **cyanobacteria** (once thought to be algae and called “blue-green algae”) about 3 billion years ago, the first organisms capable of carrying out photosynthesis and producing oxygen, which for them was a waste product. This raised the oxygen content of the atmosphere from virtually zero to the current value of 21% (by volume of dry air). The result was conversion of the atmosphere to an oxidizing medium. Vast deposits of solid iron minerals now used for iron ore were formed when atmospheric oxygen reacted with dissolved Fe in the oceans to produce solid iron oxide.

$$4\,Fe^{2+} + O_2 + 4\,H_2O \rightarrow 2\,Fe_2O_3 + 8\,H^+$$

Part of the oxygen generated by photosynthesis dissolved in water, where it was available for the development of organisms that used oxygen to metabolize organic matter. Whereas Earth’s surface had been a most inhospitable place for the existence of life, the oxygen released by photosynthesis enabled the formation of the ultraviolet-radiation-filtering layer of ozone (O_3) in the stratosphere that made life possible outside the protective confines of water. Thus life became possible on Earth’s land surface, soil was formed, aided by the weathering action of organisms including cyanobacteria that grew on rock surfaces, plants growing in soil became well established, and animals developed. The huge changes made possible by the action of single-celled cyanobacteria carrying out photosynthesis are obvious.

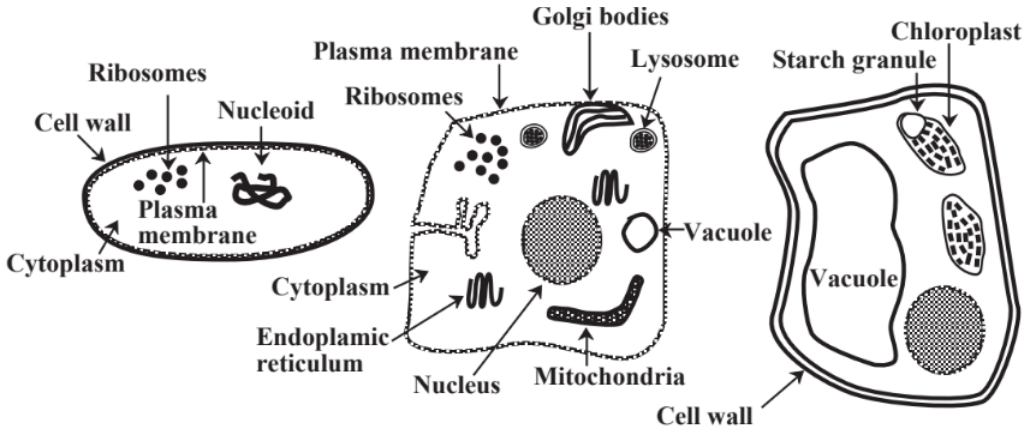


Figure 12.1

Representations of prokaryotic cells (left), eukaryotic animal cells (center), and eukaryotic plant cells (right). Eukaryotic cells are very complex and not all of the many organelles are shown

Prokaryotic cells characteristic of bacteria are enclosed by strong **cell walls** composed largely of carbohydrates that hold the cells together. The plasma membrane controls passage of materials into and out of the cell and is the site of photosynthesis in photosynthetic bacteria. Gelatinous **cytoplasm** composed largely of protein and water fills the cell. There is not a defined nucleus, but the cell has a mass of genetic material (DNA) that composes a **nucleoid**. The DNA directs cell metabolism and reproduction. Proteins are made in the cell in **ribosomes** that are distributed around the cell interior. Ribosomes and other bodies in the prokaryotic cell are not enclosed by separate defined membranes as is the case with more complex eukaryotic cells.

Major Features of Eukaryotic Cells

Animal and plant cells shown in Figure 12.1 represent the two major kinds of **eukaryotic cells** that compose all organisms other than bacteria and cyanobacteria. The major features of eukaryotic cells include the following:

- **Cell membrane**, which encloses the cell and determines what enters and leaves the cell interior. The cell membrane has varying **permeability** for various substances so that one of its crucial functions is regulation of the passage of ions, nutrients, lipid-soluble (“fat-soluble”) substances, metabolic products, toxicants, and toxicant metabolites into and out of the cell interior thus protecting the contents of the cell from undesirable outside influences. One of the adverse effects of some toxicants is damage to the cell membrane causing the cell to function improperly.
- **Cell nucleus**, which controls cell function and the genetic material required for reproduction. **Deoxyribonucleic** acid (**DNA**) discussed in Section 7.6 is the key substance in the nucleus. Damage to DNA by foreign substances may cause mutations, cancer, birth defects, defective immune system function, and other toxiceffects.
- **Cytoplasm** composed of a water-soluble proteinaceous filler called **cytosol** fills the interior of the cell not occupied by the nucleus or other bodies. Bodies of **cellular organelles**, such as mitochondria or chloroplasts are suspended in the cytoplasm.
- **Mitochondria** mediate energy conversion and utilization where carbohydrates, proteins, and fats are broken down to yield carbon dioxide, water, and energy, which is then used by the cell. The best example of this is the oxidation of the sugar glucose, $C_6H_{12}O_6$ in a process called **cellular respiration** (see Reaction 12.4.1).
- **Ribosomes** are involved in protein synthesis.
- **Endoplasmic reticulum** is the site of enzymatic metabolism of some toxicants.
- **Lysosome**, a type of organelle that contains potent substances capable of hydrolyzing and breaking down food material that enters the cell through a **food vacuole**.
- **Golgi bodies** are flattened bodies of material in some types of cells that serve to hold and release substances produced by the cells.
- **Cell walls** of provide stiffness and strength in cell walls composed primarily of cellulose (see Section 7.3 and Figure 7.3).
- **Vacuoles** inside plant cells often contain materials dissolved in water.
- **Chloroplasts** in plant cells that are the sites of photosynthesis, the chemical process which uses energy from sunlight to convert carbon dioxide and water to organic matter which is stored in the chloroplasts as **starch grains**.

This page titled 2.5.12.3: Cells - Basic Units of Life is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.3: Cells - Basic Units of Life** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.4: Metabolism and Control in Organisms

As discussed in Section 7.8 in respect to the processing of biochemicals, living organisms continually process materials and energy, a process called **metabolism**. *Photosynthesis*, which is mentioned above, is the metabolic process that provides the base of the food chain for most organisms. Animals break down complex food materials to smaller molecules through the process of **digestion**. **Respiration** occurs as nutrients are metabolized to yield energy:

$$C_6H_{12}O_6\ (\text{glucose}) + 6\,O_2\ \rightarrow 6\,CO_2 + 6\,H_2O + \text{energy}$$

Organisms assemble small molecules to produce biomolecules, such as proteins, by a **synthesis** process.

In addition to viewing metabolism as a phenomenon within an individual organism, it can be viewed as occurring within groups of organisms living in an ecosystem. Consider, for example, the metabolism of nitrogen within an ecosystem. Elemental nitrogen from the atmosphere may be fixed as organic nitrogen by bacteria living symbiotically on the roots of leguminous plants, then converted to nitrate when the nitrogen-containing biomass decays. The nitrate may be taken up by other plants and incorporated into protein. The protein may be ingested by animals and the nitrogen excreted as urea in their urine to undergo biological decay and return to the atmosphere as elemental nitrogen. Carbon from carbon dioxide in the atmosphere may be incorporated into biomass by plant photosynthesis, then eventually returned to the atmosphere as carbon dioxide as the biomass is used as a food source by animals.

Enzymes in Metabolism

In Chapter 5, Section 5.5, *catalysts* were defined as materials that enable a reaction to occur without themselves being consumed. Living organisms have catalysts that are very important in metabolism. These catalysts are special proteins that enable biochemical reactions to take place called **enzymes**. Enzymes speed up metabolic reactions by as much as almost a billion-fold. In addition to making reactions go much more rapidly, enzymes are often highly specific in the reactions that they catalyze. The reason for the specificity of enzymes is that they have very specific structures that fit with the substances upon which they act.

Enzymes were discussed in Chapter 7, Section 7.7, and their action illustrated in Figure 7.9 with respect to their processing of biochemicals. The first step in the function of enzymes is the reversible formation of an enzyme/substrate complex that forms because of the complementary shapes of the enzyme (more specifically the active site on the enzyme) and the substrate. The second step is the formation of products accompanied by release of the unchanged enzyme molecule. A very common enzymatic process called **hydrolysis** involves splitting a molecule accompanied by the addition of water with an H atom going to one of the products and an OH group to the other. Other types of enzyme-catalyzed reactions occur, including the joining of two molecules, modifications of organic functional groups on substrate molecules, and rearranging the structures of molecules.

The names of enzymes, usually ending in "*-ase*" often reflect their functions and may also indicate where they operate. An example is *gastric proteinase*, a name that indicates the enzyme acts in the stomach (*gastric*) and hydrolyzes proteins (*proteinase*). The enzyme released by the pancreas that hydrolyzes fats is called *pancreatic lipase.*

A number of factors can affect enzyme action. One important factor is temperature. Organisms without temperature-regulating mechanisms have enzymes that increase in activity as temperature increases up to the point where the heat damages the enzyme, after which the activity declines precipitously with increasing temperature. Enzymes in mammals function optimally at body temperature (37°C for humans) and are permanently destroyed by about 60°C. There is particular interest in enzymes that function in bacteria that live in hot springs and other thermal areas where the water is at or near boiling. These enzymes may turn out to be very useful in commercial biosynthesis operations where the higher temperature enables reactions to occur faster. Acid concentration also affects enzymes, such as those that function well in the acidic environment of the stomach, but stop working when discharged into the slightly basic environment of the small intestine (were this not the case, they would tend to digest the intestine walls).

A significant concern with potentially toxic substances is their adverse effects upon enzymes. As an example, organophosphate compounds, such as insecticidal parathion and military poison sarin "nerve gas" bind with acetylcholinesterase required for nerve function, causing it not to act and stopping proper nerve action. Some substances cause the intricately wound protein structures of enzymes to come apart, a process called denaturation, which stops enzyme action. The active sites of enzymes at which substrates are recognized have a high population of -SH groups. Heavy metals, such as lead and cadmium, have a strong affinity for−SH groups and may bind at enzyme active sites thus destroying the function of the enzymes.

Enzymes are of significant concern in the practice of green chemistry. One obvious relationship is that between enzymes and chemicals that are toxic to them. In carrying out green chemical processes, such chemicals should be avoided wherever possible. Another obvious relationship has to do with the use of biological processes to perform chemical operations, which are usually done under much milder and environmentally friendly conditions biologically than chemically. Biochemical processes are all carried out by enzymes. For example, several enzymes, starting with hexokinase, are involved in the multistepped biochemical fermentation synthesis of ethyl alcohol from carbohydrate glucose. With recombinant DNA technology (see Section 12.8) it is now possible to invest bacteria with enzyme systems from other organisms designed to carry out desired biochemical processes. Bacteria are much more amenable to handling and usually much more efficient than the organisms from which the genes for the desired enzyme systems are taken. Another approach is to use isolated enzymes immobilized on a solid support to carry out biochemical processes without the direct involvement of an organism.

Nutrients

The raw materials that organisms require for their metabolism are **nutrients**. Those required in larger quantities include oxygen, hydrogen, carbon, nitrogen, phosphorus, sulfur, potassium, calcium, and magnesium and are called **macronutrients**. Plants and other autotrophic organisms use these nutrients in the form of simple inorganic species, such as H_2O and CO_2, which they obtain from soil, water, and the atmosphere. Heterotrophic organisms obtain much of the macronutrients that they need as carbohydrates, proteins and lipids (see Chapter 7) from organic food material.

An important consideration in plant nutrition is the provision of **fertilizers** consisting of sources of nutrient nitrogen, phosphorus, and potassium. A large segment of the chemical manufacturing industry is involved with fixing nitrogen from the atmosphere as ammonia, NH_3, and converting it to nitrate (NO_3^-), urea (CON_2H_4), or other compounds that are applied to the soil as nitrogen fertilizer. Phosphorus is mined as mineral phosphate that is converted to biologically available phosphate ($H_2PO_4^-$ and HPO_4^{2-} ions) by treatment with sulfuric or phosphoric acid. Potassium is mined as potassium salts and applied directly as fertilizer. The ongoing depletion of sources of phosphorus and potassium fertilizer is a sustainability issue of significant concern.

Organisms also require very low levels of a number of **micronutrients**, which are usually used by essential enzymes that enable metabolic reactions to occur. For plants, essential micronutrients include the elements boron, chlorine, copper, iron, manganese, sodium, vanadium, and zinc. The bacteria that fix atmospheric nitrogen required by plants require trace levels of molybdenum. Animals require in their diet elemental micronutrients including iron and selenium as well as micronutrient vitamins consisting of small organic molecules.

Control in Organisms

Organisms must be carefully regulated and controlled in order to function properly. A major function of these regulatory functions is the maintenance of the organism's homeostasis, its crucial internal environment. The most obvious means of control in animals is through the **nervous system** in which messages are conducted very rapidly to various parts of the animal as **nerve impulses**. More advanced animals have a brain and spinal cord that function as a **centralnervous system** (CNS). This sophisticated system receives, processes, and sends nerve impulses that regulate the behavior and function of the animal. Effects on the nervous system are always a concern with toxic substances. For example, exposure to organic solvents that dissolve some of the protective lipids around nerve fibers can lead to a condition in which limbs do not function properly called **peripheral neuropathy**. Therefore, a major objective of green chemistry is to limit the use of and human exposure to such solvents.

Both animals and plants employ **molecularmessengers** that move from one part of the organism to another to carry messages by which regulation occurs. Messages sent by these means are much slower than those conveyed by nerve impulses. Molecular messengers are often **hormones** discussed as lipids in Section 7.5. Hormones are carried by a fluid medium in the organism, such as the bloodstream, to cells where they bind to **receptorproteins** causing some sort of desired response. For example, the process may cause the cell to synthesize a protein to counteract an imbalance in homeostasis. Some hormones called **pheromones** carry messages from one organism to another. They commonly serve as sex attractants. Some biological means of pest control use sex pheromones to cause sexual confusion in pesticidal insects, thus preventing their reproduction. Figure 12.2 shows a common plant hormone and a common animal hormone.

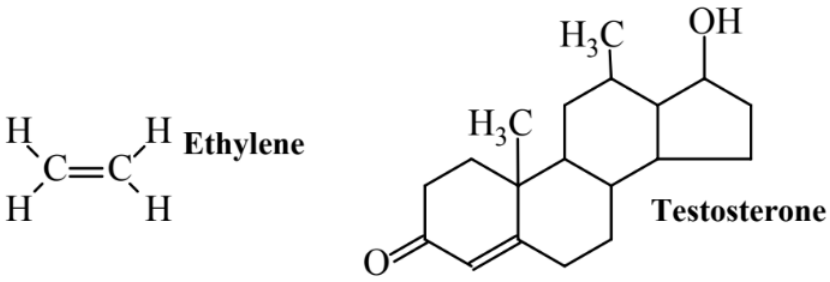


Figure 12.2. A simple molecule that acts as a plant hormone to promote maturation processes (ethylene) and a common animal hormone, testosterone, the male sex hormone. See Chapter 6, Section 6.3, for the meaning of line formulas, such as those shown for testosterone.

In animals, regulatory hormones are commonly released by **endocrineglands** shown for humans in Chapter 7, Figure 7.6. Endocrine glands in humans include the anterior pituitary gland that releases human growth hormone, the parathyroid gland that releases a hormone to stimulate uptake of calcium into the blood from bones and the digestive tract, and the pancreas that release insulin to stimulate glucose uptake from blood. These hormones are carried to target cells in fluids external to the cells. A significant concern with toxic substances is their potential to interfere with the function of endocrine glands. Another concern is that some toxic substances may mimic the action of hormones. For example, evidence exists to suggest that premature sexual development in some young female children can be caused by ingestion of synthetic chemicals that mimic the action of the female sex hormone estrogen.

This page titled 2.5.12.4: Metabolism and Control in Organisms is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.4: Metabolism and Control in Organisms** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.5: Reproduction and Inherited Traits

As noted in the preceding section, one of the major activities of organisms is metabolism by which organisms process materials and energy. The other major activity of all organisms is reproduction. Most organisms are capable of reproducing a large excess of their species because throughout time predators and hostile conditions have required large numbers of juveniles to ensure survival of enough members to continue the species. Unrestrained reproduction, especially by humans, poses a strong threat of overpopulation that will outstrip Earth's resources and is a major concern related to reproduction and the environment. A second major concern is the potential effect of environmental chemicals upon reproduction and the threat of such chemicals to cause birth defects. Therefore, chemicals that may affect reproduction are given strong consideration in the practice of green chemistry.

Primitive single-celled organisms, particularly bacteria, undergo **asexual** reproduction in which a cell simply splits to form two cells. Humans and most other multicelled organisms undergo **sexualreproduction** requiring that male sperm cells fertilize female egg cells to produce cells capable of dividing and producing new individuals.

Reproduction is directed by **genes** which occur in molecules of deoxyribonucleic acid, DNA, discussed in Chapter 7, Section 7.6. The DNA of an individual, which in sexual reproduction has contributions from both parents, determines the physical, biochemical, and behavioral traits of the organism. The DNA can be altered resulting in changes called **mutations**. A minuscule fraction of mutations are desirable and convey advantages to an individual that are passed along as heritable characteristics in offspring. This is the process of *naturalselection* that has resulted in literally millions of different species of organisms.

Some chemicals are capable of producing mutations. Control of production and exposure to these **mutagens** is a major thrust of green chemistry. This is particularly so because substances that cause mutations are generally regarded as being capable of causing cancer as well and substances that give positive tests for mutagenicity are suspect carcinogens.

This page titled 2.5.12.5: Reproduction and Inherited Traits is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.5: Reproduction and Inherited Traits** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.6: Stability and Equilibrium of the Biosphere

In order for an organism to survive and thrive, it must reach a state of stability and equilibrium with its environment. The term given to such a state is **homeostasis** ("same status"). In maintaining homeostasis, an organism must interact with its surroundings and other organisms in its surroundings and must balance flows and processing of matter (including nutrients) and energy. On an individual basis, organisms do a remarkably good job of keeping their internal levels of water, materials such as calcium in blood, and temperature within a range conducive to their well being. Mammals have developed extraordinary capabilities of homeostasis; a healthy individual maintains its internal temperature within a few tenths of a degree. The concept of homeostasis applies to entire groups of organisms living together in ecosystems and, ultimately, to the entire biosphere. Therefore, a major objective of environmental science, including the practice of green chemistry, is to maintain and enhance conditions of homeostasis in the biosphere.

Ecology describes the interaction of organisms with their surroundings and each other. An important consideration in ecology is the manner in which organisms process matter and energy. An **ecosystem** describes a segment of the environment and the organisms in it with all of the interactions and relationships that implies. An ecosystem has means of capturing energy, almost always by plants or algae that perform photosynthesis. Light, temperature, moisture, and nutrient supplies are critical aspects of an ecosystem. Ecosystems recycle essential nutrient carbon, oxygen, nitrogen, phosphorus, sulfur, and trace elements. An important part of any ecosystem is the **food chain**, or more complicated **food webs**, in which food generated by photosynthesis is utilized by different organisms at different levels. An important aspect of the food chain in respect to persistent, poorly degradable organic chemicals that are soluble in lipid (fat) tissue occurs through the sequence of animals eaten in the food chain (small creatures in water are fed upon by small fish that are eaten by large fish that are eaten by large birds). Thus, aquatic pollutants become more concentrated in lipid tissue at the top of the food chain, a process called biomagnification (see Section 12.9). An objective of the practice of green chemistry is to avoid the generation and use of chemicals capable of biomagnification in the environment.

The surroundings over a relatively large geographic area in which a group of organisms live constitute a **biome**. There are a number of different kinds of biomes. Regions near the equator may support **tropical rain forest** biomes that stay warm all of the year and in which nutrients remain largely in the organisms (rain forest soil is often notably poor in nutrients, which are mostly held in forest biomass). Temperate regions may support **temperate deciduous forests** in which the trees grow new leaves for a warm, wet summer season and shed them for cold winters. Temperate regions may also have **grasslandbiomes** in which grass grows from a tough mass of dense roots called **sod**. **Tundra** are treeless arctic regions in which during summer only a layer of wet soil thaws above a permanently frozen foundation of permafrost. Global warming is causing some profound changes in tundra biomes.

Different kinds of biomes pose a variety of environmental challenges. Some of these have come about from the conversion of biomes to cropland. Grasslands in which the sod has been broken to support wheat and other crops have proven susceptible to wind erosion, which gave rise to the catastrophic Dust Bowl that caused such great hardship on the U. S. Great Plains during the 1930s. Climate changes resulting from global warming could change the distribution of biomes, giving rise to much larger areas of hot deserts that humans might have to learn how to utilize.

Biomes in Unexpected Places

The conventional thinking in the past was that biomes would occur only in those areas where sunlight enabled conversion of inorganic carbon to biomass that could sustain a food web. It came as a surprise to marine scientists in 1977 that thousands of meters below the surface of the Pacific ocean, far too deep for light to penetrate and without significant amounts of fallout from biomass generated in surface waters, biomes existed that teemed with tubeworms, clams, and mussels. It is now known that this abundance of life is nourished by microorganisms that thrive in hot volcanic springs and that get their energy through chemosynthesis by mediating reactions of hydrogen sulfide and other substances often toxic to more familiar organisms.

A new kind of habitat was found in 1954 with the discovery of organisms, including tubeworms over two meters long that may be centuries old, that thrive on petroleum seepage on relatively cold ocean floors. These colonies, which may contain hundreds of different species, are especially abundant on the seabed of the Gulf of Mexico, where Spanish accounts from the 1500s noted oil slicks from natural petroleum leakage.

Response of Life Systems to Stress

Organisms and the ecosystems in which they exist are subject to a number of threats that can result in loss of populations and even total destruction of the system. Natural threats include drought, flooding, fire, landslide, and volcanic eruption. Humans threaten life systems with cultivation, deforestation, mining, and severe pollution. The ability of a community of organisms to resist alteration and damage from such threats, sometimes called **inertia**, depends upon several factors and provides important lessons for the survival of the human community in the face of environmental threats. One of the basic factors involved in providing resistance of a community to damage is its overall rate of photosynthesis, its **productivity**. Another important factor is **diversity** of species so that if one species is destroyed or seriously depleted, another species may take its place. **Constancy** of numbers of various organisms is desirable; wide variations in populations can be very disruptive to a biological community. Finally, **resilience** is the ability of populations to recover from large losses.

The ability of a biological system to maintain high levels of the desirable factors listed above is commonly determined by factors other than the organisms present. This is clearly true of productivity, which is a function of available moisture, suitable climate, and nutrient-rich soil. Since all organisms depend upon the availability of good food sources, diversity, constancy, and resilience tend to follow high productivity.

Relationships Among Organisms

In a healthy, diverse ecosystem, there are numerous, often complex relationships among the organisms involved. Species of organisms strongly influence each other. And organisms may greatly alter the physical portion of the system in which they live. An example of such an influence is the tough, soil-anchoring sod that develops in grassland biomes.

In most ecosystems there is a **dominant plant species** that provides a large fraction of the biomass anchoring the food chain in the ecosystem. This might be a species of grass, such as the bluestem grass that thrives in the Kansas Flint Hills grasslands. Herbivores feed upon the dominant plant species and other plants and, in turn, are eaten by carnivores. At the end of the food cycle are organisms that degrade biomass and convert it to nutrients that can nourish growth of additional plants. These organisms include earthworms that live in soil and bacteria and fungi that degrade biological material.

In a healthy ecosystem different species compete for space, light, nutrients, and moisture. In an undisturbed ecosystem the **principle of competitive exclusion** applies in which two or more potential competitors exist in ways that minimize competition for nutrients, space, and other factors required for growth. Much of agricultural chemistry is devoted to trying to regulate the competition of weeds with crop plants. Large quantities of herbicides are applied to cropland each year to kill competing weeds. In this never-ending contest, green chemistry has an important role in areas such as the synthesis of herbicides that have maximum impact on target pests with minimum impact on the environment.

Within ecosystems there are large numbers of **symbioticrelationships** between organisms which exist together to their mutual advantage. The classic case of such a relationship is that of lichen consisting of algae and fungi growing together. The fungi anchor the system to a rock surface and produce substances that slowly degrade the rock and extract nutrients from it. The algae are photosynthetic, so they produce the biomass required by the system, which is utilized in part by the fungi. Another important symbiotic relationship is that in which nitrogen-fixing bacteria grow in nodules on leguminous plant roots. The bacteria receive nutrients from the plants in exchange for chemically fixed nitrogen required for plant nutrition.

This page titled 2.5.12.6: Stability and Equilibrium of the Biosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.6: Stability and Equilibrium of the Biosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.7: DNA and the Human Genome

In Chapter 7, Section 7.6, *deoxyribonucleic acid*, *DNA*, was discussed and it was noted that this macromolecule stores and passes on genetic information that organisms need to reproduce and synthesize proteins. Recall that DNA is composed of repeating units called *nucleotides* each consisting of a molecule of the sugar 2-deoxy-β-D-ribofuranose, a phosphate ion, and one of the four nitrogen-containing bases, adenine, cytosine, guanine, and, thymine (conventionally represented by the letters A, C, G, and T, respectively). DNA is one of two *nucleic acids*, the other one of which is *ribonucleic acid*, *RNA*. Like DNA, RNA consists of repeating nucleotides but the sugar in RNA is β-D-ribofuranose and it contains uracil instead of thymine in its bases. The structural formulas of segments of DNA and RNA are shown in Figure 7.7.

The structure of DNA is a key aspect of its function, and its elucidation by Watson and Crickin 1953 was a scientific insight that set off a revolution in biology that is going on to this day. The huge DNA molecules consist of two strands counter wound with each other and held together by hydrogen bonds; a representation of this structure is shown in Figure 7.8. In this structure, the hydrogen bonds connecting complementary bases on the two strands are represented by dashed lines. Because of their structures that make hydrogen bonding possible, adenine on one strand is always hydrogen bonded to thymine on the opposite strand and guanine to cytosine. During cell division the two strands of DNA unwind and each generates a complementary strand for the DNA of each new cell.

In organisms with eukaryotic cells, DNA is divided into units associated with protein molecules called **chromosomes**. The number of these varies with the organism; humans have 23pairs of chromosomes, a total of 46. The strands of DNA in chromosomes, in turn, are divided into sequences of nucleotides, each distinguished by the nitrogen-containing base in it. These sequences of nucleotides give directions for the synthesis of a specific kind of protein or polypeptide. (Proteins are the biological molecules that make up much of the structure of cells and that perform most of the key functions of living organisms. Polypeptide is a general term for polymers of amino acids; proteins are the relatively long-chain polypeptides.) These specific groups of nucleotides, each of which has a specific function, are called **genes**. When a particular protein is made, DNA produces a nucleic acid segment designated *m*RNA, which goes out into the cell and causes the protein to be formed through a process called **transcription** and **translation** (the gene is said to be **expressed**).

As the units that give the directions for protein synthesis, genes are obviously of the utmost importance in living organisms. As discussed in Section 12.8, genes can now be transferred between different kinds of organisms and will direct the synthesis of the protein for which they are designed in the recipient organism. It is now known that a number of human diseases are the result of defective genes, and there is a genetic tendency toward getting other kinds of diseases. For example, certain gene characteristics are involved in susceptibility to breast cancer.

Because of the known relationship of gene characteristics to disease, the decision was made in the mid-1980s to map all the genes in the human body. This collective body of genes is called the **humangenome** and the project to map it is called the Human Genome Project. The original impetus for this project in the U. S. arose because of interest in the damage to human DNA by radiation, such as that from nuclear weapons. But, from the beginning, it was recognized that the project had enormous commercial potential, especially in the pharmaceutical industry, and could be very valuable in human health.

The announcement in 2001 that mapping of the human genome was complete promised great progress in biology, especially in medicine. Genes function by directing the synthesis of specific proteins and the action of most pharmaceuticals is to alter the activities of proteins, the drug's target. In some cases proteins are made more active and in others their activity is diminished. Knowledge of the human genome enables a better understanding of protein activity and should facilitate the development of more specific drugs, something that has developed only slowly. For example, the gene responsible for cystic fibrosis was discovered in 1989 by examination of family histories of the disease and only in 2010 were two drugs designed to combat that disease in clinical trials.

Genome Sequencing and Green Chemistry

The Human Genome Project and related genome sequencing of other organisms have a number of implications for green chemistry. One of the key goals of green chemistry is to use chemicals that have maximum effectiveness for their stated purpose with minimum side effects. This certainly applies to pharmaceuticals in which a knowledge of the human genome may enable

development of drugs that do exactly what they are supposed to do without affecting non target systems. This means that drugs can be made very efficiently with little waste material.

Some of the most important effects of DNA sequencing as it relates to green chemistry has to do with a wide variety of organisms other than humans. With an exact knowledge of DNA and the genes that it contains, it is possible to deal with organisms on a highly scientific basis in areas such as pest control and the biosynthesis of raw materials. An accurate map of the genetic makeup of insects, for example, should result in the synthesis of precisely targeted insecticides which kill target pests without affecting other organisms. Such insecticides should be effective at very low doses, thus minimizing the amount of insecticide that has to be synthesized and applied, consistent with the goals of green chemistry.

An exact knowledge of the genomes of organisms is extremely helpful in the practice of genetic engineering in which genes are transferred between species to enable production of desired proteins and to give organisms desirable characteristics, such as pest resistance. A number of medically useful proteins and polypeptides are now produced by genetically engineered microorganisms, most commonly genetically modified *Escherichiacoli* bacteria. Perhaps the greatest success with this technology has been the biosynthesis of human insulin, a lack of which causes diabetes in humans. Two genes are required to make this relatively short polypeptide which consists of only 51 amino acids. Other medically useful substances produced by genetically engineered organisms include human growth hormone, tissue plasmogen activator that dissolves blood clots formed in heart attacks and strokes, and various vaccine proteins to inoculate against diseases such as meningitis, hepatitis B, and influenza. Genetic engineering is discussed in more detail in Section 12.8.

This page titled 2.5.12.7: DNA and the Human Genome is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.7: DNA and the Human Genome** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.8: Genetic Engineering

Recombinant DNA and Genetic Engineering

Ever since humans started raising crops for food and fiber (and later animals) they have modified the genetic makeup of the organisms that they use. This is particularly evident in the cultivation of domestic corn which is physically not at all like its wild ancestors. Until now, breeding has been a slow process. Starting with domestication of wild species, selection and controlled breeding have been used to provide desired properties, such as higher yield, heat and drought tolerance, cold resistance, and resistance to microbial or insect pests. For some domesticated species these changes have occurred over thousands of years. During the 1900s, increased understanding of genetics greatly accelerated the process of breeding different varieties. The development of high-yielding varieties of wheat and rice during the "green revolution" of the 1950s has prevented (or at least postponed) starvation of millions of people. A technology that enabled a quantum leap in productivity of domestic crops was the development of **hydrids** from crossing of two distinct lines of the same crop, dating in a practical sense from the mid-1900s.

Traditional breeding normally takes a long time and depends largely upon random mutations to generate desirable characteristics. One of its greatest limitations has been that it is essentially confined to the same species, whereas more often that not, desired characteristics occur in species other than those being bred. Since about the 1970s, however, humans have developed the ability to alter DNA so that organisms synthesize proteins and perform other metabolic feats that would otherwise be impossible. Such alteration of DNA is commonly known as **genetic engineering** and **recombinant DNA** technology. Organisms produced by recombinant DNA techniques that contain DNA from other organisms are called **transgenic organisms**. With recombinant DNA technology, segments of DNA that contain information for the specific syntheses of particular proteins are transferred between organisms. Most often the recipient organisms are bacteria, which can be reproduced (cloned) over many orders of magnitude from a cell that has acquired the desired qualities. Therefore, to synthesize a particular substance such as human insulin or growth hormone, the required genetic information can be transferred from a human source to bacterial cells, which then produce the substance as part of their metabolic processes.

The mechanics of recombinant DNA gene manipulation is a complex and sophisticated operation. The first step involves lyzing (opening up) a cell that has the genetic material needed and removal of this material from the cell. Through enzyme action the sought-after genes are cut from the donor DNA chain. These are next spliced into small DNA molecules. These molecules, called **cloning vehicles**, are capable of penetrating the host cell and becoming incorporated into its genetic material. The modified host cell is then reproduced many times and carries out the desired biosynthesis.

Recombinant DNA technology is a rapidly growing area that is having profound effects, especially in agriculture and medicine. It is being used increasingly to produce crops with unique characteristics, to synthesize pharmaceuticals, and to make a variety of useful raw materials as renewable feedstocks. Recombinant DNA technology has a lot of potential in the development of green chemistry and sustainability, such as in the sustainable production of chemical feedstocks and products of various kinds. An example is synthesis of polylactic acid using lactic acid produced enzymatically with corn and polymerized by standard chemical processes. In the environmental area genetic engineering offers the potential for the production of bacteria engineered to safely destroy troublesome wastes and to produce biological substitutes for environmentally damaging synthetic pesticide.

Early concerns about the potential of genetic engineering to produce "monster organisms" or new and horrible diseases have been largely allayed, although not entirely so, and resistance to the application of recombinant DNA technology is strong in some quarters, particularly in Europe. However, caution is still required with this technology. One example of a problem has been the emergence of weeds resistant to the widely used herbicide glyphosate as discussed at the beginning of this chapter.

Once plants containing desired transgenes have been produced, an exhaustive evaluation process occurs. This process has several objectives. The most obvious of these is an evaluation of the transplanted gene's activity to see if it produces adequate quantities of the protein for which it is designed. Another important characteristic is whether or not the gene is passed on reliably to the plant's progeny through successive generations. It is also important to determine whether the modified plant grows and yields well and if the quality of its products are high.

Only a few strains of plants are amenable to the insertion of transgenes, and normally their direct descendants do not have desired productivity or other characteristics required for a commercial crop. Therefore, transgenic crops are crossbred with high-yielding varieties. The objective is to develop a cross that retains the transgene while having desired characteristics of a commercially viable

crop. The improved variety is subjected to exhaustive performance tests in greenhouses and fields for several years and in a number of locations. Finally, large numbers of genetically identical plants are grown to produce seed for commercial use.

Many kinds of genetically modified plants have been developed and more are being marketed commercially every year. These are discussed in more detail in Chapter 14.

This page titled 2.5.12.8: Genetic Engineering is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.8: Genetic Engineering** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.9: Biological Interaction with Environmental Chemicals

Organisms in the environment interact significantly with xenobiotic materials (those foreign to living systems) in their surroundings. The uptake of such materials by organisms is discussed in this section. The biodegradation of xenobiotic substances, primarily through the action of bacteria, is discussed in Section 12.10.

Bioaccumulation is the term given to the uptake and concentration of xenobiotic materials by living organisms. The materials may be present in water in streams or bodies of water, sediments in bodies of water, drinking water, soil, food, or even the atmosphere. Bioaccumulation can lead to **biomagnification** in which xenobiotic substances become successively more concentrated in the tissues of organisms higher in the food chain. This usually occurs with poorly degradable, lipid-soluble organic compounds. Suppose, for example, that such a compound contacts lake water, accumulates in solid detritus in the water, sinks to the sediment, is eaten by small burrowing creatures in the sediment, which are eaten by small fish. The small fish may be eaten by larger fish, which in turn are consumed as food by birds. At each step, the xenobiotic substance may become more concentrated in the organism and may reach harmful concentrations in the birds at the top of the food chain. This is basically what happened with DDT, which almost caused the extinction of eagles and hawks.

Fish that bioaccumulate poorly degradable, lipid-soluble organic compounds from water will lose them back to water if they are placed in an unpolluted environment. The process by which this occurs is called **depuration**. The time required to lose half of the bioaccumulated xenobiotic material is called the **half-life** of the substance.

The most straightforward case of bioaccumulation is **bioconcentration**, which occurs when a substance dissolved in water enters the body of a fish or other aquatic organism by passive processes (basically, just “dissolves” in the organism), and is carried to bodies of lipid in the organism in the blood flow. The model of bioconcentration assumes that the organism taking up the compound does not metabolize the compound, a good assumption for refractory organic compounds such as DDT or PCBs. It also assumes that uptake is by nondietary routes, including diffusion through the skin and especially through the gills of fish. The model of bioconcentration applies especially to substances that have low water solubilities (though high enough to make the compound available for uptake) and high lipid solubilities. This model of bioconcentration assumes a dynamic equilibrium between the xenobiotic substance dissolved in water and the same substance dissolved in lipid tissue. It is called the **hydrophobicitymodel** because of the hydrophobic (“water-hating”) nature of the substance being taken up.

The degree of bioconcentration depends upon a number of factors. The most important of these are the relative water and lipid solubility of the compound. The size and shape of the xenobiotic molecule also seem to be factors, as is temperature. In addition, bioconcentration depends upon the species of fish and their age, size, and lipid contents. Bioconcentration may be expressed by **bioconcentrationfactors** defined as

$$\text{Bioconcentration factor} = \frac{\text{Concentration of xenobiotic in lipid}}{\text{Concentration of xenobiotic in water}}$$

The bioconcentration factor can also be regarded as the ratio of the solubility of the compound in lipid to its solubility in water. Typical bioconcentration factors for PCBs and hexachlorobenzene in sunfish, trout, and minnows range from somewhat more than 1,000 to around 50,000, reflecting the high lipid solubility of these compounds.

This page titled 2.5.12.9: Biological Interaction with Environmental Chemicals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.9: Biological Interaction with Environmental Chemicals** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.10: Biodegradation

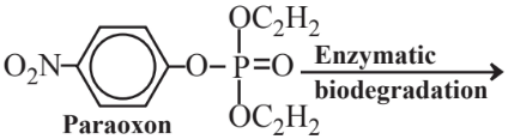
Bacteria, fungi, and protozoa in the environment play an important role in biodegrading both natural materials and synthetic substances. These processes occur predominantly in water, in sediments in bodies of water, and in soil. Biodegradation is the process by which biomass from deceased organisms is broken down to simple inorganic constituents, thus completing the cycle in which biomass is produced from atmospheric carbon dioxide and from water by photosynthesis.

The biodegradation of substances in the environment by the action of enzymes in microorganisms can be divided metabolically into two categories. The first of these is the utilization by microorganisms of organic matter that can be metabolized for energy and as material to synthesize additional biomass. This is the route taken by microorganims degrading biomass from other organisms, and to a lesser extent in the biodegradation of some xenobiotic materials.The second way in which microorganisms metabolize environmental chemicals is through **cometabolism** in which the organism's enzymes act upon the substances as a "side-line" of their normal metabolic processes. The substances that are cometabolized are called secondary substrates because they are not the main compounds for which the enzymatic processes are designed.

A commonly cited example of cometabolism occurs with the action of *Phanerochaete chrysosporium* on organochlorine compounds, including PCBs and dioxins. Commonly known as the white rot fungus, this organism has an enzyme system that normally breaks down lignin, the degradation-resistant "glue" that holds cellulose together in wood and woody plants. Under certain stressed conditions, however, the enzyme will act to cometabolize synthetic organochlorine compounds and was once widely promoted as a means of remediating hazardous waste sites contaminated with such compounds.

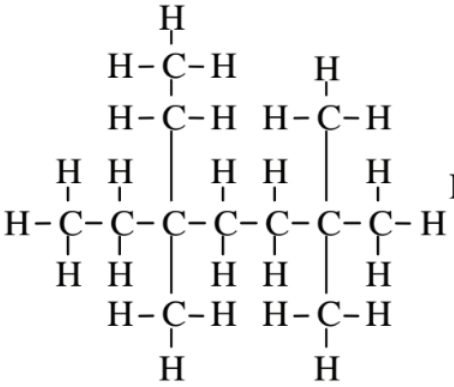
The degree of biodegradation varies over a wide range. In the simplest case, the change to the substrate molecule is relatively minor, such as addition, deletion, or modification of a functional group. Complete biodegradation to simple inorganic species—CO_2 for carbon, NH_4^+ or NO_3^- for nitrogen, HPO_4^{2-} for phosphorus, SO_4^{2-} for sulfur—is the process of **mineralization**, which is crucial in completing elemental cycles in the environment.

An important step in biodegradation is the modification of a substance to reduce its toxicity. This process is called **detoxication** (or often detoxification). An example of detoxication is given in Reaction 12.10.1 below for the conversion of insecticidal paraoxon, a potent nerve poison, to *p*-nitrophenol, which is only about 0.005× as toxic. In some cases, however, action of microorganisms in the environment may produce a much more toxic material. An example of this is the generation of highly toxic, mobile methylmercury species, $Hg(CH_3)_2$ and $HgCH_3^+$ from insoluble, relatively harmless inorganic mercury species.

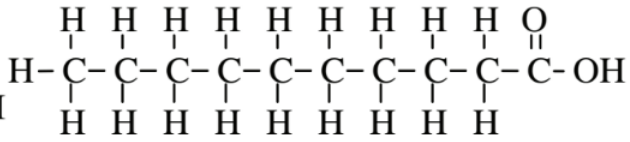


$$O_2N\text{-}C_6H_4\text{-}OH + HPO_4^{2-}, \text{ other products} \qquad (12.10.1)$$

A number of factors are involved in determining the effectiveness and rate of biodegradation.The compound in question has to be biodegradable. Biodegradability is influenced by both physical properties, such as water solubility, and chemical characteristics including the presence of functional groups amenable to microbial attack. As illustrated by the example, below,

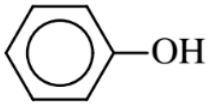


Highly branched, poorly biodegradable compound



Highly biodegradable compound with a straight hydrocarbon chain and a functional group amenable to biological attack

branched-chain hydrocarbons are very resistant to biodegradation whereas straight-chain hydrocarbons, especially those with a suitable functional group, are readily metabolized by microorganisms. It should be noted that even very poorly biodegradable compounds can often be degraded under suitable conditions. As an example, phenol,



Phenol

is a biocidal compound that kills bacteria and was once the most commonly used disinfectant. However, in dilute solution and under the appropriate conditions, phenol can be destroyed by bacteria. An important aspect of biodegradation of resistant compounds is to use microorganisms acclimated to the particular kind of compounds. Populations of acclimated microorganisms are commonly found in locations where the kinds of compounds to be treated have been spilled, such as in petroleum spills on soil.

Biodegradability of compounds is an important consideration in green chemistry. This is especially true of “consumable” materials that are dissipated to the environment. One of the most common examples of the use of a biodegradable material as a consumable material is the substitution in household detergents of biodegradable LAS surfactant, which has a readily biodegradable straight hydrocarbon chain as part of its molecular structure, in place of ABS surfactant, which has a poorly biodegradable branched chain.

This page titled 2.5.12.10: Biodegradation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.10: Biodegradation** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture

The most important use of soil and the biosphere for humans is **agriculture**, the production of food and fiber by growing crops and livestock. Agriculture is very closely tied with the practice of green chemistry in many ways. Agricultural chemicals, including fertilizers, herbicides, and insecticides are produced and applied to crops and land in enormous quantities. Annual production of millions of kilograms of these chemicals demands the proper practice of green chemistry and engineering. The judicious use of relatively small quantities of herbicides enables planting of crops in soil covered with residues of the previous year's crops with little or no cultivation of soil. This practice of low-tillage agriculture, now called conservation tillage (see Section 11.11), is in keeping with the best practice of green chemistry and industrial ecology. Organic biomass produced by plants can be used as a renewable source of raw material and fuel. Some plants are now being genetically engineered to produce specific chemicals.

The practice of agriculture is absolutely essential for the survival of humankind. In order to continue to feed growing world populations while maintaining and even enhancing the ability of soil to produce food, it is necessary that the practice of agriculture be as green as possible. In the past and still today, this has often not been true. Cultivation of soil by humans has displaced native plants, destroyed wildlife habitat, contaminated soil with pesticides, filled rivers and bodies of water with sediments, and otherwise perturbed and damaged the environment. Agricultural practices arguably represent the greatest incursion of the anthrosphere into the other environmental spheres. On the positive side, growth of domestic crops temporarily removes greenhouse gas carbon dioxide from the atmosphere and provides organic raw materials and biomass fuel without any net addition of carbon dioxide to the atmosphere.

The basis of agriculture is the development of domestic plants from their wild ancestors. (The same can be said of animals, but only a handful of animal species have been domesticated, although each consists of many different breeds.) Our prehistoric ancestors learned to select plants with desired characteristics for the production of food and fiber, developing new species that often require the careful efforts of expert botanists to relate them to their wild ancestors. Only around 1900 were the scientific principles of heredity applied to plant breeding, eventually with excellent results. Using scientific methods, agriculturists accomplished a "green revolution" in the 1950s and 1960s that resulted in varieties of rice and wheat, especially, that had vastly increased yields. The techniques used included selective breeding, hybridization, cross-pollination, and back-crossing to develop grain varieties which, combined with chemical fertilizers and pesticides, lead to much higher crop yields. India, a country on the verge of starvation in the 1940s, increased its grain output by 50%. Developments such as higher yielding and faster maturing dwarf varieties of rice enabled better nutrition for an increasing world population, at least postponing the inevitable problems that will result from population growth. By breeding plants resistant to cold, drought, and insects, overall crop productivity has been further increased. Increased nutritional values for grain have been achieved, such as the development of corn varieties that have higher levels of lysine amino acid.

One of the major advances in plant breeding has been the development of **hybrids** produced by crossing **true-breeding** strains of plants. So-called "hybrid vigor" is well known, and many hybrids have vastly greater yields than their parent strains. Corn, a remarkably productive photosynthesizer, has proven most amenable to the production of hybrids, in part because of the separation of male flowers which grow on the tops of plants from female flowers attached to the budding corn ears. By planting rows of corn that alternate between two different strains and cutting the tassels from the tops of the plants that are to produce the corn seed, hybrid corn varieties are readily produced. More recently, techniques have been developed for growing hybrids of other kinds of plants.

There are, of course, many factors other than the genetic strains of plants that are involved in high crop productivity. The effects of weather have been mitigated by the development of crop varieties that resist heat, cold, and drought. The provision of water by artificial irrigation has greatly increased crop productivity and is essential for crop productivity in some regions, such as the vegetable-growing areas of California. Irrigation practices continue to become more efficient with the replacement of wasteful spray irrigators by systems that apply water directly to soil, or even directly to plant roots. Computerized control of irrigation can make it much more efficient. Environmentally, widespread use of herbicides has had some excellent benefits, along with some harmful effects, by enabling greater crop productivity with less tillage of land.

This page titled 2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.11: Production of Food and Fiber by the Biosphere - Agriculture** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.12: Agricultural Applications of Genetically Modified Organisms

As discussed in Section 7.6 of Chapter 7, genes composed of deoxyribonucleic acid, DNA, located in the nuclei of cells direct cell reproduction and synthesis of proteins and generally direct the organism activities. Plant scientists are now able to modify DNA by processes called *recombinant DNA technology*. (This technology is also being applied to a lesser extent to animals.) Recombinant DNA technology normally involves taking a single characteristic from one organism— the ability to produce a bacterially synthesized insecticide, for example — and splicing it into another organism. By so doing, for example, corn and cotton have been genetically engineered to produce their own insecticide. Plants produced by this method are called **transgenic** plants. During the 1970s, the ability to manipulate DNA through genetic engineering became a reality, and during the 1990s, it became the basis of a major industry. This technology promises some exciting developments in agriculture and, indeed, is leading to a "second green revolution." Direct manipulation of DNA can greatly accelerate the process of plant breeding to give plants that aremuch more productive, resistant to disease, and tolerant to adverse conditions. In the future, entirely new kinds of plants may even be engineered.

Plants are particularly amenable to recombinant DNA manipulation. In part this is because huge numbers of plant cells can be grown in appropriate media and mutants can be selected from billions of cells that have desired properties, such as virus resistance. Individual plant cells are capable of generating whole plants, so cells with desired qualities can be selected and allowed to grow into plants which may have the qualities desired. Ideally, this accomplishes in weeks what conventional plant breeding techniques would require decades to do.

Transgenic crops have many detractors, and demonstrations have broken out and test plots of crops destroyed by people opposed to what they call "Frankenfoods." Opposition has been especially strong in Europe, and the European Commission, the executive body of the European Union, has disallowed a number of transgenic crops. Despite these concerns, transgenic crops are growing in importance and they have become widely utilized in highly populated countries, particularly China, where they are seen as a means of feeding very large populations.

The Major Transgenic Crops and their Characteristics

The two characteristics most commonly developed in transgenic crops is tolerance for herbicides that kill competing weeds and resistance to pests, especially insects, but including microbial pests (viruses) as well. In the earlier years of transgenic crop plantings most of the crops had traits for only one of these characteristics, but in more recent years so-called stacked varieties with two or more characteristics have become more common and now comprise the fastest growing sector of transgenic crops. As of 2008, the land areas planted to transgenic crops in the eight leading countries were the following (millions of hectares in parentheses): United States (62.5), Argentina (21.0), Brazil (15.8), India (7.6), Canada (7.6), China (3.8), Paraguay (2.7), and South Africa (1.8 million hectares). The most common biotech crops are the following (hectares planted in 2008 in parentheses): Soybeans (65.8), maize (corn, 37.5), cotton (15.5), and canola (5.9). Herbicide tolerance has been the predominant biotech trait with 79 million hectares (of a global total of 125 million hectares), next were stacked traits (27 million hectares), then insect resistance (19 million hectares).

In 2010 Monsanto and Dow Agrosciences introduced stacked transgenic corn with 8 traits including resistance to insects above and below the soil as well as tolerance of some common herbicides. It is claimed that this variety will reduce the refuge area for corn planting from 20% to 5%. (The refuge is a fraction of the area of a crop that is planted to non-transgenic crops to generate enough insecticide-susceptible insects to dilute the resistant ones that eventually develop in the transgenic areas. The rationale for this approach is that insects growing in refuge areas without any incentive to develop resistance will crossbreed with resistant strains, preventing them from becoming dominant.) In addition to pest resistance and tolerance to herbicides, future stacked transgenic crops are expected to have characteristics such as drought resistance, high omega-3 lipid production in soybeans and elevated levels of pro-Vitamin A in Golden Rice.

The disruption of natural ecosystems by cultivation of land and planting agricultural crops provides an excellent opportunity for opportunistic plants — weeds — to grow in competition with the desired crops. To combat weeds, farmers use large quantities of a variety of herbicides. The heavy use of herbicides poses a set of challenging problems. In many cases, to be effective without causing undue environmental damage, herbicides must be applied in specified ways and at particular times. Collateral damage to crop plants, environmental harm, and poor biodegradation leading to accumulation of herbicide residues and contamination of

water supplies are all problems with herbicides. A number of these problems can be diminished by planting transgenic crops that are resistant to particular herbicides discussed above. The most common such plants are those resistant to Monsanto's Roundup herbicide (glyphosate, structural formula shown in Section 12.1).

This widely used compound is a broad-spectrum herbicide, meaning that it kills most plants that it contacts. One of its advantages from an environmental standpoint is that it rapidly breaks down to harmless products in soil, minimizing its environmental impact and problems with residue carry-over. By using "Roundup Ready" crops, of which by far the most common are transgenic soybeans, the herbicide can be applied directly to the crop, killing competing weeds. Application when the crop plants are relatively small, but after weeds have had a chance to start growing, kills weeds and enables the crop to get a head start. After the crop has developed significant size, it deters the growth of competing weeds by shade that deprives the weeds of sunlight.

Aside from weeds, the other major class of pests that afflict crops consists of a variety of insects. Two of the most harmful of these are the European corn borer and the cotton bollworm, which cost millions of dollars in damage and control measures each year and can even threaten an entire year's crop production. Even before transgenic crops were available, *Bacillus thuringiensis* (Bt) was used to control insects. This soil-dwelling bacterium produces a protein called delta-endotoxin. Ingested by insects, delta-endotoxin partially digests the intestinal walls of insects causing ion imbalance, paralyzing the system, and eventually killing the insects. Fortunately, the toxin does not affect mammals or birds. Bt has been a popular insecticide because as a natural product it degrades readily and has gained the acceptance often accorded to "natural" materials(many of which are deadly).

Genetic engineering techniques have enabled transplanting genes into field crops that produce Bt. This is an ideal circumstance in that the crop being protected is generating its own insecticide, and the insecticide is not spread over a wide area. There are several varieties of insecticidal Bt, each produced by a unique gene. Several insecticidal pests are well controlled by transgenic Bt. In addition to the European corn borer mentioned above, these include the Southwestern corn borer and corn earworm. Cotton varieties that produce Bt are resistant to cotton bollworm. Bt-producing tobacco resists the tobacco budworm. Potato varieties have been developed that produce Bt to kill the Colorado potato beetle, although this crop has been limited because of concerns regarding Bt in the potato product consumed directly by humans. Although human digestive systems are not affected adversely by Bt, there is concern over its being an allergen because of its proteinaceous nature.

Virus resistance in transgenic crops has concentrated on papaya. This tropical fruit is an excellent source of Vitamins A and C and is an important nutritional plant in tropical regions. The papaya ringspot virus is a devastating pest for papaya, and transgenic varieties resistant to this virus are now grown in Hawaii. One concern with virus-resistant transgenic crops is the possibility of transfer of genes responsible for the resistance to wild relatives of the plants that are regarded as weeds, but are now kept in check by the viruses. For example, it is possible that virus-resistant genes in transgenic squash may transfer to competing gourds, which would crowd out the squash grown for food.

Future Crops

The early years of transgenic crops can be rather well summarized by soybeans, corn, and cotton resistant to herbicides and insects. In retrospect, these crops will almost certainly seem rather crude and unsophisticated. In part, this lack of sophistication is due to the fact that the genes producing the desired qualities are largely expressed by all tissues of the plants and throughout their growth cycle, giving rise to problems such as the Bt-contaminated corn pollen that may threaten Monarch butterflies or Bt-containing potatoes that may not be suitable for human consumption. It is anticipated that increasingly sophisticated techniques will overcome these kinds of problems and will lead to much improved crop varieties in the future.

There are many potential green chemistry aspects from genetic engineering of agricultural crops. One promising possibility is to increase the efficiency of photosynthesis, which is only a few tenths of a percent in most plants. Doubling this efficiency should be possible with recombinant DNA techniques, which might significantly increase the production of food and biomass by plants. For example, with some of the more productive plant species, such as fast-growing hybrid poplar trees and sugarcane, biomass is competitive with fossil fuels as an energy source. A genetically engineered increase in photosynthesis efficiency could enable biomass to economically replace expensive petroleum and natural gas for fuel and raw material. A second possibility with genetic engineering is the development of the ability to support nitrogen-fixing bacteria on plant roots in plants that cannot do so now. If corn, rice, wheat, and cotton could be developed with this capability it could save enormous amounts of energy and natural gas (a source of elemental hydrogen) now consumed to make ammonia synthetically.

A wide range of other transgenic crops are under development. One widely publicized crop is“golden rice” which incorporates β-carotene in the grain, which is therefore yellow, rather than the normal white color of rice. The human body processesβ-carotene to Vitamin A, the lack of which impairs vision and increases susceptibility to maladies including respiratory diseases, measles, and diarrhea. Since rice is the main diet staple in many Asian countries, the widespread distribution of golden rice could substantially improve health. As an example of the intricacies of transgenic crops, two of the genes used to breed golden rice were taken from daffodil and one from a bacterium! Some investigators contend that humans cannot consume enough of this rice to provide a significant amount of Vitamin A.

As of 2010, transgenic alfalfa and sugar beets resistant to glyphosate herbicide were being promoted for agricultural use. Alfalfa is a nutritious forage crop for animal feed, a legume that grows synergistically with nitrogen-fixing *Rhizobium* bacteria growing in nodules attached to its roots. In 2010 the U.S. Supreme Court overturned a lower court decision that had prevented widespread distribution of these crops because of the possibility that their glyphosate-resistant qualities might spread to other plants and violate restrictions on foods designated as “organic” and that some countries have put in place against all transgenic food.

Work continues on improved transgenic oilseed crops. The one getting the most attention is canola, the source of canola oil. Efforts are underway to modify the distribution of oils in canola to improve the nutritional value of the oil. Another possibility is increased Vitamin E content in transgenic canola. Sunflower, another source of vegetable oils, is the subject of research designed to produce improved transgenic varieties. Herbicide tolerance and resistance to white mold are among the properties that are being developed in transgenic sunflowers.

Decaffeinated coffee and tea have become important beverages. Unfortunately, the processes that remove caffeine from coffee beans and tea leaves also remove flavor, and some such processes use organic solvents that may leave undesirable residues. The genes that produce caffeine in coffee and tea leaves have now been identified, and it is possible that they may be removed or turned off in the plants to produce coffee beans and tea leaves that give full-flavored products without the caffeine. Additional efforts are underway to genetically engineer coffee trees in which all the beans ripen at once, thereby eliminating the multiple harvests that are now required because of the beans ripening at different times.

Although turf grass for lawns would not be regarded as an essential crop, enormous amounts of water and fertilizers are consumed in maintaining lawns and grass on golf courses and other locations. Healthy grass certainly contributes to the “green” esthetics of a community. Furthermore, herbicides, insecticides, and fungicides applied to turf grass leave residues that can be environmentally harmful. So the development of improved transgenic varieties of grass and other groundcover crops can be quite useful. There are many desirable properties that can benefit grass. Included are tolerances for adverse conditions of water and temperature, especially resistance to heat and drought. Disease and insect resistance are desirable. Reduced growth rates can mean less mowing, saving energy. For grass used on waterways constructed to drain excess rain runoff from terraced areas (see Figure 11.6) a tough, erosion-resistant sod composed of masses of grass roots is very desirable. Research is underway to breed transgenic varieties of grass with some of these properties. Also, grass is being genetically engineered for immunity to the effects of Roundup herbicide (see above), which is environmentally more benign than some of the herbicides such as 2,4-D currently used on grass.

An interesting possibility for transgenic foods is to produce foods that contain vaccines against disease. This is possible because genes produce proteins that resemble the proteins in infectious agents, causing the body to produce antibodies to such agents. Diseases for which such vaccines may be possible include cholera, hepatitis B, and various kinds of diarrhea. The leading candidate as a carrier for such vaccines is the banana. This is because children generally like bananas and this fruit is readily grown in some of the tropical regions where the need for vaccines is the greatest.

This page titled 2.5.12.12: Agricultural Applications of Genetically Modified Organisms is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.12: Agricultural Applications of Genetically Modified Organisms** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.13: The Anthrosphere in Support of the Biosphere

Although humans are commonly blamed — usually with ample justification — for adverse effects upon the biosphere, human activities and modern technology have a high potential for benefitting the biosphere. This is especially true now that there is such an explosion in the understanding of biological sciences. The realization that the biosphere can contribute immeasurably to the benefit of humankind through such things as the provision of renewable feedstocks as raw materials provides a strong incentive to use technology to the benefit of the biosphere.

The most direct interface between the biosphere and technology occurs in agriculture. The production of biomass per unit area of land has increased in a spectacular fashion in recent decades with the use of fertilizers, herbicides, insecticides, and sophisticated means of cultivation and harvesting. Now the application of recombinant DNA technology (see Section 12.12) to agriculture promises even greater advances. In the past, the ways in which techniques for improved agricultural productivity were applied were largely divorced from considerations of the natural ways in which plants and animals grow on land. Fortunately, there is a growing realization of the important information that nature can provide in maintaining agricultural productivity. For example, in the prevention of water erosion, terraces constructed on land are designed to funnel excess water runoff onto grassed waterways. By planting these waterways to native grasses, a tough, erosion-resistant sod can be established that stands up under the punishment of occasional deluges of runoff water while surviving intermittent severe droughts. On a larger scale, in place of cultivating drought-prone prairie land to grow grain to feed to cattle, a better approach may be to reseed these lands to tough native grasses and allow bison to feed upon the grass as a source of meat (less fat and more healthy than beef from cattle).

The restoration and development of “natural” areas has become an important endeavor commonly termed **restoration ecology**. This often is advised with farmland that is too marginal to support profitable agricultural operations. The example of restoring native grasslands was mentioned above. Much of the rocky, hilly, unproductive farmland in New England is now reverting to forests. In such restoration efforts, modern construction machinery with the capacity to move enormous quantities of dirt have proven useful. One example in which such machinery is used is in leveling large areas for the construction of wetlands. Rivers that were once straightened to facilitate water flow — with catastrophic results in the form of flooding and erosion — are now being restored with the bends and meanders that characterize a healthy river. As discussed in Chapter 8, Section 8.2, following the catastrophic 500-year floods on the Missouri and Mississippi Rivers in 1993, large areas of cropland in the river bottoms were purchased by the Federal Government, river dikes designed to prevent flooding were breached, and the land was allowed to revert to a wild state. Land disrupted by strip mining has been smoothed over to reduce erosion, topsoil applied, and trees planted to produce natural areas and wildlife habitat.

A significant amount of restoration ecology has been devoted to restoring game animals, some of which had been driven virtually to extinction by over-hunting and habitat destruction. Animals that have come back in significant numbers include wild turkeys, wood ducks, snowy egrets, and American bison. Some of these efforts have been almost too successful. Once endangered Canadian geese have greatly increased in numbers and now populate many suburban areas where they often show their displeasure with sharing their new habitats with humans by hissing, flapping their wings aggressively, and even pinching exposed flesh with their sturdy beaks. In many areas deer now destroy crops and are a traffic hazard. Sophisticated captive breeding techniques are now used to reproduce endangered species of animals, and animal cloning may reach a point at which these efforts are routine.

In the area of green chemistry, sophisticated chemical analysis techniques can now be used to find and eliminate the sources of chemical hazards to wildlife. The classic example of this occurred in the 1960s when it was found that insecticidal DDT, biomagnified through the food chain, was preventing reproduction of endangered eagles and hawks at the top of the chain. In 1970 a newly developed technique for the determination of mercury showed that large fish were contaminated by this heavy metal released from sediments by bacterial methylation. Analysis of lipid tissue in humans, caribou, and polar bears now indicate a global distillation mechanism by which persistent organic compounds evaporate into the atmosphere in warmer regions of Earth and condense in the polar regions, leading to significant contamination of food supplies. One of the major objectives of green chemistry is the elimination of the generation and use of such materials.

As the projected effects of global warming become more pronounced during the next century, technology will be employed to a greater extent to deal with these effects upon the biosphere. Increasingly sophisticated genetic engineering techniques will be used to develop plant varieties that can withstand the heat and drought resulting from global warming. Another possibility is the

development of plants that can grow in saltwater. Using renewable solar and wind energy, vast water desalination projects will be developed to provide fresh water to irrigate high-value crops where the costs can be justified.

This page titled 2.5.12.13: The Anthrosphere in Support of the Biosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.13: The Anthrosphere in Support of the Biosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.14: Livestock and their Wastes

Areas where agriculture is practiced intensively with heavily fertilized crops and large populations of livestock can cause serious water pollution problems. This is true of the watershed of Chesapeake Bay, the largest U.S. estuary with a watershed that includes areas of New York, Pennsylvania, Delaware, Maryland, Virginia, and West Virginia. One of the agricultural areas that discharges runoff to the bay is Lancaster County, Pennsylvania, that is the location of more than 5000 farms with some of the most productive soil in the world. More than half of these farms are owned by Amish, a religious sect that does not use electricity, automobiles, and other modern conveniences. These farms are livestock-intensive with heavy production of dairy products. Unlike most modern farms that use tractors and other farm equipment powered by diesel and gasoline engines, the Amish rely upon horses and mules for cultivation. The result of these factors is that each year Lancaster County produces about 9 million kilograms of manure plus large quantities urine, more than any of the other counties in the region. Both of these materials are rich in nitrogen, phosphorus, and potassium that have resulted in drainage into Chesapeake Baycontributing to heavy growths of algae leading to eutrophication. As a result of algal nutrient pollution Chesapeake Bay has had a dead zone that has been the subject of remediation efforts since the 1970s. In addition to runoff pollution, problems have been reported of pollution of farm wells in the county contaminated with *Escherichiacoli* bacteria, which are indicators of animal feces pollution and some wells have shown elevated levels of nitrate ion that can come from the biodegradation of nitrogenous biomaterials. (Nitrate is generally harmless to adult humans but in human infants and ruminant animals including cattle can cause methemoglobinemia, a condition in which the iron(II) in blood hemoglobin is oxidized to iron(III) producing methemoglobin, which does not transport blood in the bloodstream.)

Attempts are underway to reduce the water pollution from livestock in Lancaster and other livestock-intensive counties by measures including minimal tillage agriculture, cover crops, reduction and containment of barnyard runoff, forested buffers along streams, and manure pits to collect wastes from concentrated livestock feeding operations. Farms such as those in Lancaster County actually provide a great opportunity to practice sustainable, green agriculture. Horses and mules used for cultivation require no petroleum, the animal wastes provide abundant fertilizer, and anoxic biodegradation of animal wastes can generate large quantities of methane gas, a fuel for engines and household heating that is probably acceptable to religious sects who reject modern, especially electrically-powered, devices. A farm operated with such sustainable practices could serve as a largely self-sufficient industrial ecosystem (an alternative term that does not include“industrial” would probably be more acceptable to the operators).

This page titled 2.5.12.14: Livestock and their Wastes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **12.14: Livestock and their Wastes** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.12.15: Questions and Problems

1. Fill in the blanks in the following statement: Living organisms carry out chemical processes in _____, which are bodies of the order of a _____ in size.

2. Name several constituents of living cells. What are two constituents typical of plant cells?

3. What are hormones? Where are they produced?

4. To what general class of kinds of biomolecules do enzymes belong? What is their function?

5. How are enzymes named?

6. How does temperature relate to enzyme action, particularly as it pertains to enzymes acting in the body?

7. What is recombinant DNA? Why is it produced?

8. What is metabolism? How is respiration related to metabolism?

9. What can natural ecosystems teach humans about how to manage their industrial systems?

10. Distinguish between bioaccumulation and biomagnification as they apply to lipid-soluble organic compounds.

11. What are six characteristics shared by all living organisms?

12. What are the four basic kinds of molecules that comprise living organisms?

13. Photosynthesis is the reaction that provides food and energy to keep biological systems going. What important biological process is the opposite of photosynthesis?

14. What are the six kingdoms of living organisms?

15. Justify the statement that human beings are oxic heterotrophs.

16. What is the general effect of good living conditions, such as mild climate and abundant nutrients, on ecosystems?

17. Give four organelles found in eukaryotic cells.

18. What huge effect did photosynthesis by cyanobacteria have upon the atmosphere in the early years of life on earth?

19. In order to use the food that humans eat, it must undergo digestion and respiration. What is the distinction between these two phenomena?

20. What is the explanation of high levels of persistent organic compounds in nonindustrialized polar regions?

21. Describe the overall process of enzyme action.

22. Name two important aspects of enzymes as they relate to green chemistry.

23. Which three elements are commonly regarded as plant fertilizers?

24. Name two ways in which messages are sent in the body of an animal. Do they have about the same speed?

25. Give examples of organs that produce hormones.

26. Describe a concern pertaining to synthetic chemicals and hormones.

27. Name a concern other than mutations regarding exposure to mutagens.

28. In order for an organism to live and thrive, its internal environment must be maintained within acceptable limits. Give the name of this state.

29. Name some of the critical factors in an ecosystem. How do functioning ecosystems handle materials?

30. Tropical rain forests, temperate deciduous forests, and grasslands are all examples of_____.

31. Name four characteristics that enable life systems to respond well to stress.

32. What are symbiotic relationships between organisms? Give an example.

33. Name several factors that have contributed to the increased agricultural productivity of biomass and another factor that promises such increases in the future.

34. Define the human genome and give reasons why it is so important.

35. What are the relationships among cell DNA, RNA, and ribosomes in protein synthesis?

36. Name some substances that are beneficial to human health that are made by genetically engineered organisms. Name the organism often used for this purpose.

37. What was the “green revolution?” What technology will likely result in a “second green revolution?”

38. How is transgenic technology distinguished from conventional breeding techniques? What kind of biomolecule makes transgenic technology possible?

39. Distinguish between bioaccumulation and biomagnification

40. What are xenobiotic substances? What kind of metabolic process is largely responsible for their biodegradation?

2.5.12.15: Questions and Problems is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** has no license indicated.

2.5.12.16: Supplementary References

Brand, Stewart, *Whole Earth Discipline: Why Dense Cities, Nuclear Power, Transgenic Crops, Restored Wildlands, and Geoengineering are Necessary*, Penguin, New York, 2010.

Brooker, Robert, Eric Widmaier, Linda Graham, and Peter Stiling, *Biology*, McGraw Hill, New York, 2010.

Dotretsov, Nicolay, *Biosphere Origin and Evolution*, Springer, Berlin, 2007.

Herons, Two, *The Biosphere: Protecting our Global Environment*, 4th ed., Kendall Hunt Publishing, DuBuque, IA, 2008.

Ikerd, John E., *Crisis and Opportunity: Sustainability in American Agriculture (Our Sustainable Future*, Bison Books, Winnipeg, 2008.

Kole, Chittaranjan, Charles Michler, Albert G. Abbott, and Timothy C. Hall, *Transgenic Crop Plants*, Springer, Berlin, 2010.

Raven, Peter H., Linda R. Berg, and David M. Hassenzahl, *Environment*, 6th ed., Wiley, Hoboken, NJ, 2008.

Manning, Richard, *Against the Grain: How Agriculture has Hijacked Civilization*, North Point Press, New York, 2005.

Mazoyer, Marcel, and Laurence Roudart, *A History of World Agriculture: From the Neolithic Age to the Current Crisis*, New York, 2006

Smil, Vaclav, *The Earth's Biosphere: Evolution, Dynamics, and Change*, MIT Press, Cambridge, MA 2003.

2.5.12.16: Supplementary References is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Supplementary References** has no license indicated.

SECTION OVERVIEW

2.5.13: The Anthrosphere, Industrial Ecology, and Green Chemistry

"We are addicted to growth. That addiction to growth stokes the greed that drives the endless and often pointless consumption that we have defined as economic success. The problem with being addicted to growth is that we live on a finite planet. No matter what growth's apologists claim about finding more resources or harnessing new technology, an addiction to growth, by definition, must at some point collide with reality" (Rudy Baum, *Chemical and Engineering* News, *June 28, 2010*).

2.5.13.1: Industrial Ecology and Industrial Ecosystems

2.5.13.2: Metabolic Processes in Industrial Ecosystems

2.5.13.3: Life Cycles in Industrial Ecosystems

2.5.13.4: Kinds of Products

2.5.13.5: Attributes Required by an Industrial Ecosystem

2.5.13.6: The Kalundborg Industrial Ecosystem

2.5.13.7: Environmental Impacts of Industrial Ecosystems

2.5.13.8: Green Chemistry and Industrial Ecology

2.5.13.9: Predicting and Reducing Hazards

2.5.13.10: The E-Factor in Green Chemistry

2.5.13.11: Catalysts and Catalysis

2.5.13.12: Biocatalysis with Enzymes

2.5.13.13: Energizing Chemical Reactions and Process Intensification

2.5.13.14: Solvents and Alternate Reaction Media

2.5.13.15: Feedstocks and Reagents

2.5.13.16: Literature Cited and Supplementary References

2.5.13.17: Questions and Problems

This page titled 2.5.13: The Anthrosphere, Industrial Ecology, and Green Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.13.1: Industrial Ecology and Industrial Ecosystems

Recall that the *anthrosphere* has been defined as a fifth sphere of the environment, the one made and modified by human activities. As such, it has developed in ways that are often in conflict with other spheres of the environment, including even the human denizens of the biosphere who have constructed the anthrosphere. This has given rise to the many environmental, resource, and sustainability problems that afflict the world today.

It is crucial for humankind and, indeed, the Earth as a whole, that the anthrosphere be brought into a state of compatibility with the other environmental spheres and with Earth. Green chemistry has a key role to play in this endeavor. In a sense, green chemistry is all about the greening of the anthrosphere. In order to understand how this may occur, it is necessary to introduce and explain the key concept of **industrial ecology**. Industrial ecology integrates the principles of science, engineering, and ecology in industrial systems through which goods and services are provided in away that minimizes environmental impact and optimizes utilization of resources, energy, and capital. In so doing, industrial ecology considers every aspect of the provision of goods and services from concept, through production, and to the final fate of products remaining after they have been used. Industrial ecology considers industrial systems in a closed-loop model rather than a linear one thereby emulating natural biological ecosystems, which are sustainable by nature. Industrial ecology is above all a **sustainable** means of providing goods and services.

Industrial ecology works through groups of industrial concerns, distributors, and other enterprises functioning to mutual advantage, using each others' products, recycling each others' potential waste materials, and utilizing energy as efficiently as possible. By analogy with natural ecosystems, such a system is an **industrial ecosystem**. Successful industrial ecosystems achieve the maximum possible degree of recycling. To quote Kumar Patel of the University of California at Los Angles, "The goal is *cradle to reincarnation*, since if one is practicing industrial ecology correctly there is no grave." As has been the case with natural ecosystems, the best means of assembling industrial ecosystems is through natural selection in which the various interests involved work out mutually advantageous relationships. However, with a knowledge of the feasibility of such systems, external input and various kinds of incentives can be applied to facilitate the establishment of industrial ecosystems. A key measure of the success of such a system can be given by the following relationship:

$$\frac{\text{Market value of products}}{\text{Consumption of material and energy}}$$

Just as organisms in natural ecosystems develop strong symbiotic relationships — the inseparable union of algae and fungi in lichens growing on rock surfaces, for example — concerns operating in industrial ecosystems develop a high degree of **industrial symbiosis**. It is the development of such mutually advantageous interactions between two or more industrial enterprises that result in the self-assembly of an industrial ecosystem in the first place. The recycling components of an industrial ecosystem are absolutely dependent upon symbiotic relationships with their sources of supply.

Figure 13.1 outlines a general industrial ecosystem. The major inputs to such a system are energy and virgin raw materials. A successful system minimizes use of virgin raw materials and maximizes efficiency of energy utilization. The materials processing sector produces processed materials such as sheet steel or synthetic organic polymers. These in turn go to a goods fabrication sector in which the processed materials are formed and assembled or, in the case of consumables such as detergents, formulated to give the desired product. Scrap materials, rejected

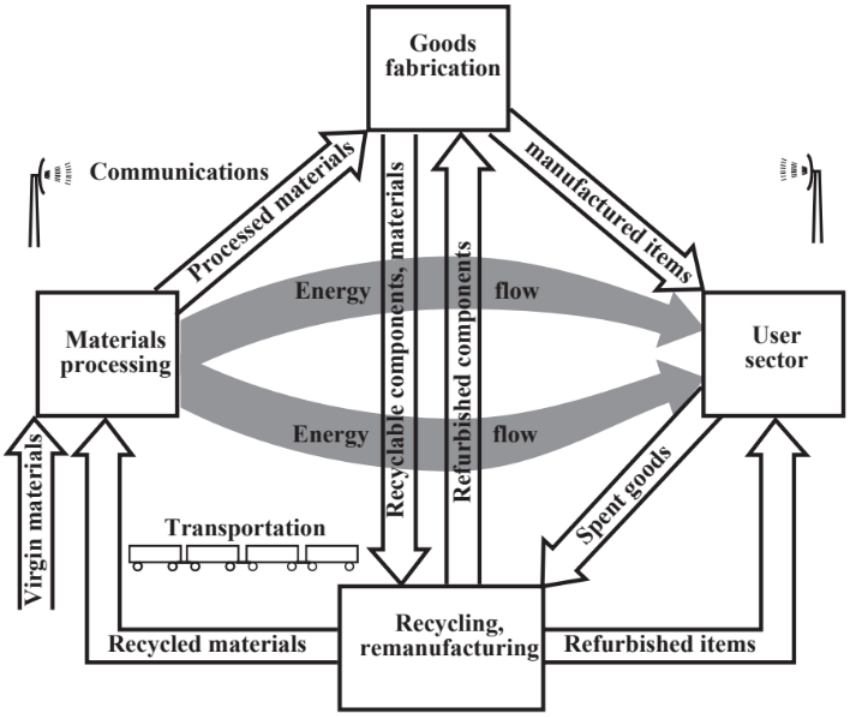


Figure 13.1. Outline of the major components of an industrial ecosystem.

components, and off-specification consumables generated during goods fabrication may go to recycling and remanufacturing. From goods fabrication, manufactured items or formulated substances are taken to a user sector, which includes consumers and industrial users. In a successful system of industrial ecology, waste materials from the user sector are minimized and, ideally, totally eliminated. Spent goods from the user sector are taken to recycling and remanufacturing to be introduced back into the materials flow of the system. Such items may consist of automobile components that are cleaned, have bearings replaced, and otherwise refurbished for the rebuilt automobile parts market. Another typical item is paper, which is converted back to pulp that is made into paper again. In some cases the recycling and remanufacturing sector salvages materials that go back to materials processing to start the whole cycle over. An example of such a material is scrap aluminum that is melted down and recast into aluminum for goods fabrication. Communications are essential to a successful industrial ecosystem, as is a reliable, rapid transportation system. It is especially important that these two sectors work well in modern manufacturing practice which calls for “just in time” delivery of materials and components to avoid the costs of storing such items.

An important characteristic of an industrial ecosystem is its **scope**. A regional scope large enough to encompass several industrial enterprises, but small enough for them to interact with each other on a constant basis is probably the most satisfactory scale to consider. Frequently such systems are based around transportation systems. Segments of interstate highways over which goods and materials move between enterprises by truck may constitute industrial ecosystems.

This page titled 2.5.13.1: Industrial Ecology and Industrial Ecosystems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.1: Industrial Ecology and Industrial Ecosystems** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.2: Metabolic Processes in Industrial Ecosystems

Industrial metabolism refers to the processes to which materials and components are subjected in industrial ecosystems. It is analogous to the metabolic processes that occur with food and nutrients in biological systems. Like biological metabolism, industrial metabolism may be addressed at several levels. A level of industrial metabolism at which green chemistry, especially, comes into play is at the molecular level where substances are changed chemically to give desired materials or to generate energy. Industrial metabolism can be addressed within individual unit processes in a factory, at the factory level, at the industrial ecosystem level, and even globally.

A significant difference between industrial metabolism as it is now practiced and natural metabolic processes relates to the wastes that these systems generate. Natural ecosystems have developed such that true wastes are virtually nonexistent. For example, even those parts of plants that remain after biodegradation of plant materials form soil humus (see “Soil Organic Matter” in Chapter 11, Section 11.8) that improves the conditions of soil on which plants grow. Anthropogenic industrial systems, however, have developed in ways that generate large quantities of wastes, where a waste may be defined as *dissipative use of natural resources*. Furthermore, human use of materials has a tendency to dilute and dissipate materials and disperse them to the environment. Materials may end up in a physical or chemical form from which reclamation becomes impractical because of the energy and effort required. A successful industrial ecosystem overcomes such tendencies.

Organisms performing their metabolic processes degrade materials to extract energy(catabolism) and synthesize new substances (anabolism). Industrial ecosystems perform analogous functions. The objective of industrial metabolism in a successful industrial ecosystem is to make desired goods with the least amount of byproduct and waste. This can pose a significant challenge. For example, to produce lead from lead ore for the manufacture of lead/acid storage batteries requires mining large quantities of ore, extracting the relatively small fraction of the ore consisting of lead sulfide mineral, and roasting and reducing the mineral to get lead metal. The whole process generates large quantities of lead-contaminated tailings left over from mineral extraction and significant quantities of byproduct sulfur dioxide, which must be reclaimed to make sulfuric acid and not released to the environment. The recycling pathway, by way of contrast, takes essentially pure lead from recycled batteries and simply melts it down to produce lead for new batteries; the advantages of recycling in this case are obvious.

There are some interesting comparisons between natural ecosystems and industrial systems as they now operate. The basic unit of a natural ecosystem is the organism, whereas that of an industrial system is the firm or, in the case of large corporations, the branch of a firm. Natural ecosystems handle materials in closed loops; with current practice, materials traverse an essentially one-way path through industrial systems. It follows that natural systems completely recycle materials, whereas in industrial systems the level of recycling is often very low. Organisms have a tendency to concentrate materials. For example, carbon in carbon dioxide that is only about 0.04% of atmospheric air becomes concentrated in organic carbon through photosynthesis. Industrial systems in contrast tend to dilute materials to a level where they cannot be economically recycled, but still have the potential to pollute. Aside from maintaining themselves during their limited lifetime, the major function of organisms is reproduction. Industrial enterprises do not have reproduction of themselves as a primary objective; their main function is to generate goods and services in a manner that maximizes monetary income.

Unlike natural ecosystems in which reservoirs of needed materials are essentially constant(oxygen, carbon dioxide, and nitrogen from air as examples) industrial systems are faced with largely depleting reservoirs of materials. For example, the lead ore cited above is a depleting resource; more may be found, but only a finite amount is ultimately available. Fossil energy resources are also finite. For example, much more fossil energy from coal may be available, but its utilization as the world’s main source of energy over the long term would come at an unacceptable cost of global warming from carbon dioxide emissions. Again, industrial metabolic processes that emphasize recycling are desirable because recycling gives essentially constant reservoirs of materials in the recycling loop. Ideally, even in the case of energy, renewable energy resources such as wind and solar power provide an essentially constant, non-depleting energy source.

As discussed under “Control in Organisms” under Section 12.4 of Chapter 12, biological ystems have elaborate systems of control. Considering the metabolism that occurs in an entire natural ecosystem, it is **self-regulating**. If herbivores that consume plant biomass become too abundant and diminish the stock of the biomass, their numbers cannot be sustained, the population dies back, and their food source rebounds. The most successful ecosystems are those in which this self-regulating mechanism operates continuously without wide variations in populations. Industrial systems do not inherently operate in a self-regulating manner that is

advantageous to their surroundings, or even to themselves in the long run. Examples of the failure of self-regulation of industrial systems abound in which enterprises have wastefully produced large quantities of goods of marginal value, running through limited resources in a short time, and dissipating materials to their surroundings, polluting the environment in the process. Despite these bad experiences, within a proper framework of laws and regulations designed to avoid wastes and excess, industrial ecosystems can be designed to operate in a self-regulating manner. Such self-regulation operates best under conditions of maximum recycling in which the system is not dependent upon a depleting resource of raw materials or energy.

Obviously, recycling is the key to the successful function of industrial metabolism. Figure13.2 illustrates the importance of the **level of recycling**. In low-level recycling, a material or component is taken back to near the beginning of the steps through which it is made. For example, an automobile engine block might be melted down to produce molten metal from which new blocks are then cast. With high-level recycling, the item or material is recycled as close to the final product as possible. In the case of the automobile engine block, it may be cleaned, the cylinder walls rehoned, the flat surfaces replaned, and the block used as the platform for assembling a rebuilt engine. In this example and many others that can be cited, high-level recycling uses much less energy and materials and is inherently more efficient. The term given to the value attributed to an item or material recycled near the top of the energy/materials pyramid shown in Figure 13.2 is called its **embedded utility**.

This page titled 2.5.13.2: Metabolic Processes in Industrial Ecosystems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.2: Metabolic Processes in Industrial Ecosystems** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.3: Life Cycles in Industrial Ecosystems

In conventional industrial systems, a product is manufactured and marketed after which the vendor forgets about it (unless some product defect, such as sticking accelerator pedals on an automobile forces a recall). In a system of industrial ecology, however, the entire **life cycle** of the product is considered. An important aspect of such a consideration is the **life-cycle assessment**. The overall goal of a life-cycle assessment is to determine, measure, and minimize environmental and resource impacts of products and services.

An important decision that must be made at the beginning of a life-cycle assessment is determination of the **scope** of the assessment. Parameters included in the scope include the time period to be considered, the area (space) to be considered, and the kinds of materials, processes, and products that will go into the assessment. As an example, consider the chemical synthesis of an insecticide that releases harmful vapors and generates significant quantities of waste material. A narrowly focussed life-cycle assessment might take account of control measures to capture released vapors and the best means of disposing of the waste byproducts. A broader scope would consider a different synthetic process that might not cause the problems mentioned. An even broader scope might consider whether or not the insecticide even needs to be made and used; perhaps there are more acceptable alternatives to its use.

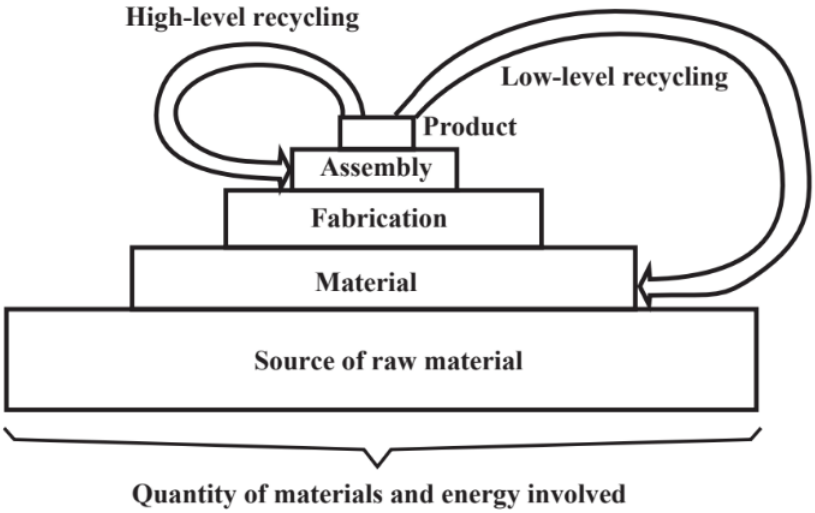


Figure 13.2. The level at which recycling occurs strongly influences the amount of energy required and the quantity of materials that must be processed.

Life cycle assessment involves an **inventoryanalysis** to provide information about the consumption of material and release of wastes from the point that raw material is obtained to make a product to the time of its ultimate fate, an **impactanalysis** that considers the environmental and other impacts of the product, and an **improvementanalysis** to determine measures that can be taken to reduce impacts. A life-cycle assessment gives a high priority to the choice of materials in a way that minimizes wastes. It considers which materials and whole components can be used or recycled. And it considers alternate pathways for manufacturing processes or, in the case of chemical manufacture, alternate synthesis routes.

In doing life-cycle assessments it is useful to consider the three major categories of **products**, **processes**, and **facilities**, all of which have environmental and resource impacts. Products are obviously the things and commodities that consumers use. They are discussed further in Section13.4. Processes refer to the ways in which products are made. Facilities consist of the infrastructural elements in which products are made and distributed.

Some of the greater environmental impacts from commerce result from the processes by which items are made. An example of this is paper manufacture. The environmental impact of the paper *product* tends to be relatively low. Even when paper is discarded improperly, it does eventually degrade without permanent effect. But the *process* of making paper, beginning with harvesting of wood and continuing through the chemically intensive pulping process and final fabrication has significant environmental impact. In addition to potential air and water pollution, paper making consumes energy and requires large amounts of water. Processes can be made much more environmentally friendly by application of the principles of industrial ecology enabling maximum recycling of materials that otherwise would have significant pollution potential.

The impact of facilities can vary over a wide range. A specialized facility such as a steel mill or petroleum refinery can have a significant environmental impact. Abandoned sites of these facilities can be blighted and difficult and expensive to restore for some other use. (The term

“brownfields” is sometimes used to describe sites of abandoned industrial facilities, and restoration of blighted “brownfields” is often a major goal of urban renewal projects.) One of the more challenging kinds of facilities to decommission are sites of nuclear power reactors in which there is a significant amount of radioactivity to deal with in dismantling and disposing of some of the reactor components. The impact of facilities can be minimized by designing them with future use and eventual decommissioning in mind. Typically, well designed commercial buildings may have a number of lives during which they are used by a sequence of enterprises. A key aspect of abuilding destined for multiple use is structure flexibility so that it can easily be rearranged for new uses.

Product Stewardship

The control of the life cycle of products that are sold is difficult. Attempts are made to exercise some degree of control by facilitating return of spent products. For example, laser printer cartridges commonly come with a return address and postage to return the used cartridge to the manufacturer. Another approach is to pay for the return of spent products. The price of a new automobile battery is commonly increased to cover a modest refund paid upon return of the spent battery. Automobile tires have a financial disincentive for recycling in that the customer is commonly charged a disposal fee.

The success rate with these kinds of measures is mixed. Automobile tires are almost invariably mounted at a dealer’s place of business, so customers pay the disposal fee in order to not be responsible for used tire disposal. Customers who install their own batteries usually turn in the old ones because they do not have a good disposal option. The return rate with printer cartridges is probably lower because people do not want to go to the trouble to mail them back.

The rate of return of items at the end of their life cycles is higher in companies and institutions where returns are more readily facilitated. The presence in a large office of a recycled paper bin or an “out box” where returned laser printer cartridges can be dropped for mailing certainly increases the recycle rate of these items. One key to recycling is in product**stewardship** where there are clear lines of accountability for items and materials. A good way to ensure product stewardship for things such as office equipment is through the leasing of equipment. In such a case, the concern holding the lease, rather than the user, is responsible for the ultimate fate of the item. The same approach can even be used with some kinds of materials, such as activated carbon used in a water treatment plant. Here the concern leasing the carbon retains ownership and is responsible for picking up the spent material for recycling. Another way to help ensure product stewardship through the consumer sector is by charging deposits on items and refunding the deposit upon return for recycling.

This page titled 2.5.13.3: Life Cycles in Industrial Ecosystems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.3: Life Cycles in Industrial Ecosystems** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.4: Kinds of Products

In considering life-cycle assessments, it is useful to divide products into three major categories. The first of these are **consumableproducts** which, by the nature of their use, are used up or dispersed to the environment with no possibility of recovery. Such materials include laundry detergents that are flushed down the drain with wash-water or windshield washer fluid, which is squirted onto windshields, then wiped off. Another class of product consists of **recyclable commodities**. Engine antifreeze and motor oil are potentially recyclable commodities in that, in principle (though somewhat rarely in practice), they can be reclaimed after use, purified, reformulated, and sold again. **Service products** (sometimes called durable products) are usually devices that have multiple uses and last for a long time. The washing machine in which consumable laundry detergent is used is a typical service product.

Since consumable products are dispersed to the environment, it is important that they have environmentally friendly characteristics. They should first of all be **nontoxic** at the levels and manner in which organisms are exposed to them. In addition to not causing acute toxicity, they should not be carcinogenic or mutagenic and should not cause birth defects. Another characteristic that consumable products should have is that they should not be **bioaccumulative**. As discussed under the topic of “Biological Interaction with Environmental Chemicals” in Section 12.9, **bioaccumulation** is the term given to the uptake and concentration of xenobiotic materials by living organisms. Poorly biodegradable, lipid-soluble materials such as PCB compounds have a strong tendency to bioaccumulate, and such substances should be avoided in consumable products. Consumable products should also be **degradable**. The most common type of degradation is biodegradation, which occurs primarily through the action of microorganisms. The practice of green chemistry can aid in making biodegradable products by, for example, avoiding branched-chain hydrocarbon structures in organic compounds and by attaching functional groups, such as the organic carboxylic acid group, $-CO_2H$, that are amenable to microbial attack.

Recyclable commodities should be designed with durability and recycling in mind. In order for them to last through a normal life cycle, such commodities should not be as degradable as consumables. An example of making a product more amenable to recycling is the use of bleachable and degradable inks on newsprint, which makes it easier to recycle the newsprint to produce a grade of recycled paper that meets acceptable color standards.

Although service products are designed to last for relatively long times, they do reach a stage requiring disposal or recycling. A key factor in recycling is the availability of channels through which such products can be recycled. Proposals have been made for “de-shopping” centers where items such as old computers and broken small appliances can be returned for recycling. Service products should be designed and constructed to facilitate disassembly so that various materials can be separated for recycling.

This page titled 2.5.13.4: Kinds of Products is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.4: Kinds of Products** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.5: Attributes Required by an Industrial Ecosystem

Under the heading of “Response of Life Systems to Stress” in section 12.6, is a discussion of the *inertia* of biological communities, which is their resistance to alteration and damage. The key factors involved in inertia were mentioned as *productivity* of basic food materials, *diversity* of species, *constancy* of numbers of various organisms and *resilience* in the ability of populations to recover from loss. Industrial ecosystems likewise have key attributes that are required for their welfare. These include **energy**, **materials**, and **diversity**. One big difference between biological and industrial ecosystems is the time scale involved. In the evolution of organisms, a time period of several thousand years is very brief, whereas in industrial systems several decades may be a very long time.

Energy

With enough energy, almost anything is possible (see Chapter 15, “Sustainable Energy: The Essential Basis of Green Systems”). Therefore, the provision of adequate amounts of energy that can be used without damaging the environment too much is essential for the function of industrial ecosystems. And the energy that is available has to be used as effectively and efficiently as possible. It was once believed that the world’s vast coal resources would provide enough energy to meet human needs for several centuries. Now it is apparent that consuming most of these energy resources would cause unacceptable global warming effects. Solar energy and wind energy, which derives from solar energy, come about as close as any energy source to offering ideal renewable sources of energy. But there are major problems with the intermittent nature of these sources and the need that they present for short-term energy storage. Furthermore, they both require vast areas of land in order to provide a significant share of energy needs. Then there are unexpected problems, such as the one arising from the accumulation of dead insects on windmill blades, spoiling their finely tuned aerodynamic characteristics and reducing power output by about half in strong winds. Properly run nuclear power facilities can provide abundant energy for many decades, but this source comes with its own set of problems and is strongly opposed by many.

Cogeneration represents the most efficient energy use within an industry or within an industrial ecosystem. The two major reasons that an industrial plant uses energy are (1) for steam used in processing, such as heating chemical reaction mixtures to cause a reaction to go faster, and(2) to generate electricity. Traditionally, industrial operations, such as petroleum refineries, have bought electricity from external power plants to run pumps and compressors, for lighting, and other purposes that consume electricity. Steam, which can only be shipped economically for relatively short distances, is normally generated by burning fossil fuels in boilers on the site. Since a maximum of only approximately 40% of the heat generated in burning a fuel in a power plant can be converted to electricity, and because of losses in electrical transmission lines, obtaining electricity from an external source is a relatively inefficient means of getting power. Much greater efficiencies can be attained by burning fuels, such as natural gas, in large turbines connected to an electrical generator and using the hot exhaust from the turbine to raise steam. This approach can double the overall efficiency of energy utilization.

Materials

There are several approaches to providing materials. These can be classified as **dematerialization** in which less material is used for a specific purpose, **substitution** of abundant materials for scarce ones, **recycling** materials, and **wastemining** in which needed materials are extracted from wastes.

Examples abound of areas in which the need for materials has been reduced in recent decades. Higher voltage electrical transmission carries more power over thinner copper or aluminum wires, the switch from 6-volt to 12-volt electrical systems in automobiles has enabled lighter wiring in automotive electrical systems, modern photographic film uses much less silver than in years past(and the almost complete switch to digital photography has virtually eliminated the use of silver in photography), and the switch from biased-ply to radial tires has greatly extended tire life, so that much less rubber is required (as well as saving fuel by lowered rolling resistance). Dematerialization has been spectacular in the electronics area. The popular laptop computer has far more computing power than did the earliest vacuum-tube-equipped computers that each required an entire air conditioned building for housing.

Material substitution is an area in which green chemistry has made a significant contribution and will continue to do so at an accelerating pace in the future. The most spectacular advances have been made in electronics where material substitution, which

enabled dematerialization to occur, has provided electronic circuits with many orders of magnitude more capability than the circuits that they replaced. The glowing, electricity-consuming vacuum tubes, capacitors, resistors, and transformers of the receiver circuit of a 1950s table-top radio have been replaced with a tiny circuit almost invisible to the human eye. The huge numbers of copper wires that carried telegraph and telephone messages in the 1940s have now been replaced by fiber optic signal conductors that carry unimaginably more information per unit mass of carrier. Polyvinylchloride (PVC) pipe, synthesized from inexhaustible world resources of chlorine and potentially from biomass hydrocarbon sources, has replaced copper and steel for water and wastewater transmissions. Toxic liquid sulfur dioxide and ammonia used in early refrigerator models were replaced by nontoxic, nonflammable chlorofluorocarbons (CFCs). When the CFCs were found to deplete stratospheric ozone, substitutes were developed containing at least one bound H atom per molecule (HCFCs) that break down in the troposphere before reaching stratospheric altitudes. Many more similar examples could be cited.

Recycling is of course one of the major objectives of a system of industrial ecology and one in which significant progress is being made. There are some consumable items that are not practical to recycle and for which the raw materials are abundant enough that recycling is not required. Household detergents are in this category. A second group of recyclables are those that are not particularly scarce, but for which recycling is feasible and desirable. Wood and paper fall into this category. A third category of recyclable materials consists of metals, particularly the more valuable and scarce ones, such as chromium, platinum, and palladium. These metals definitely should be recycled. A fourth category of recyclables consists of parts and apparatus that can be refurbished and reused.

Waste mining, the extraction of useful materials from wastes, provides more materials while benefitting the environment. One of several important examples of waste mining is the extraction of combustible methane gas, a low-polluting premium fossil fuel, from municipal refuse landfills in which the biodegradation of organic matter in the absence of oxygen generates the gas. Sulfur in sulfur dioxide extracted from the flue gases generated in burning coal that contains sulfur can be reclaimed and used to make sulfuric acid. Methods have been developed to extract aluminum from finely divided coal fly ash generated in coal combustion. In this case, the finely divided, homogeneous, dry nature of the fly ash is a definite advantage in processing it. It is anticipated that growing scarcity of resources combined with the need to dispose of a variety of wastes will lead to the development of additional waste mining processes in the future.

Diversity

Diversity in industrial ecosystems, tends to impart a **robust** character to them, which means that if one part of the system is diminished, other parts will take its place and keep the system functioning well. Many communities that have become dependent upon one or just a few major enterprises have suffered painful economic crises when a major employer leaves or cuts back. The fouling of beaches in Louisiana, Mississippi, Alabama, and Florida from the 2010 BP Deepwater Horizon has devastated the tourist trade and forced painful economic adjustments. In many parts of the world, water supply from a single vulnerable source threatens diversity.

This page titled 2.5.13.5: Attributes Required by an Industrial Ecosystem is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.5: Attributes Required by an Industrial Ecosystem** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.6: The Kalundborg Industrial Ecosystem

Industrial ecosystems of various degrees of sophistication have existed ever since the first industries were established, although they were not called industrial ecosystems or even recognized as such until the latter 1980s. The system most commonly cited as a fully developed industrial ecosystem is the one that developed spontaneously in Kalundborg, Denmark. This system is centered around two very large energy enterprises. The first of these is the ASNAES electrical plant fired by coal and with a capacity of 1,500 megawatts. The second enterprise is the Statoil petroleum refinery, processing 4–5 million tons per year. The initial relationships in the Kalundborg system consisted of exchanges between these two enterprises. The power plant sold steam to the refinery to use in processing, and the refinery provided the power plant with fuel gas and cooling water. Both enterprises produce low-level energy in the form of steam that is used for district heating of homes and commercial buildings. The heat is also used in a large greenhouse operation, as well as in a fish-farming enterprise. Another big player in the Kalundborg industrial ecosystem is the Novo Nordisk pharmaceutical plant, which receives steam from the energy suppliers. This huge enterprise makes 40% of the world's supply of insulin as well as industrial enzymes. Fermentation processes in this plant generate excess yeast, which is used as protein supplement for swine. The plant produces large quantities of biological sludge, which, along with wastewater treatment sludge from the waste and wastewater treatment plant associated with the fish farm, is used as fertilizer and soil conditioner in area farms.

The Kalundborg industrial ecosystem is often cited for the spontaneous way in which it developed, beginning in the 1960s with steam and electricity provided to the petroleum refinery from the power plant then branching out to a large variety of other enterprises in the vicinity. Some of the enterprises were driven by required measures to lower pollution. As a result of the requirement for lime scrubbing of the stack gas from the power plant, large quantities of calcium sulfate were produced, which were used to manufacture gypsum wallboard for buildings. Air pollution control measures resulted in the substitution of clean burning hydrocarbons from the petroleum refinery in place of some of the coal to generate electricity in the power plant. The requirement to remove sulfur from petroleum led to the construction of a sulfuric acid plant.

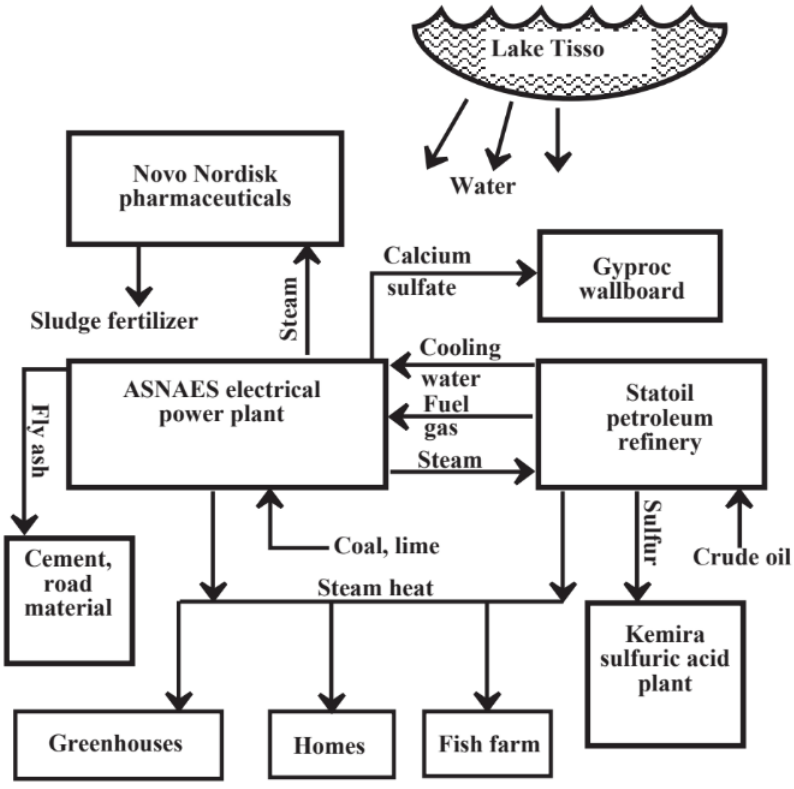


Figure 13.3. The industrial ecosystem in Kalundborg, Denmark, is commonly cited as an example of a functioning system of industrial ecology. The figure shows the numerous mutually advantageous operations in this system. It developed spontaneously beginning with exchanges of materials and energy between the electrical power plant and the petroleum refinery.

It is interesting to consider the conditions that lead to such a well developed industrial ecosystem at Kalundborg. First of all, the Kalundborg industrial ecosystem did not develop from directives from any centralized authority mandating cooperation. Instead it arose from agreements between various entities acting in their own corporate self-interests. A rather close social system that promoted contact between individuals was helpful. The relatively small geographic area involved has been helpful in enabling facile communication and the transfer of materials and energy among the various enterprises. This is especially so in that several of the main commodities involved — steam, water, waste treatment sludge — cannot be shipped economically for any great distances.

This page titled 2.5.13.6: The Kalundborg Industrial Ecosystem is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.6: The Kalundborg Industrial Ecosystem** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.7: Environmental Impacts of Industrial Ecosystems

The practice of industrial ecology in the anthrosphere certainly has profound potential effects upon the atmosphere, hydrosphere, geosphere, and biosphere. Anthrospheric influences may range from highly localized effects to global effects, such as greenhouse warming or stratospheric ozone depletion. The magnitude of the effects may be minor, or they may be catastrophic. Until relatively recently, the effects of human activities on the surrounding environment were of relatively little concern, resulting in neglect that is the cause of many of the environmental problems that exist even today. However, the proper practice of industrial ecology requires that consideration be given to the various influences that anthrospheric activities have upon the surrounding environment.

Consider the kinds of effects that industrial activities may have upon the natural environmental spheres. One of the most obvious influences is upon the atmosphere because of the emission to the atmosphere of pollutant gases, vapors from volatile compounds, and particles. Released carbon dioxide and vapors such as those of fluorinated hydrocarbons have a high potential to cause greenhouse warming. Particles obscure visibility and cause adverse health effects in people who must breathe the air in which they are contained. Chlorofluorocarbons lead to stratospheric ozone depletion and hydrocarbons and nitrogen oxides released to the atmosphere can cause formation of photochemical smog. Industrial activities often utilize large quantities of water for cooling and other purposes. Water may become polluted or warmed excessively when used for cooling (thermal pollution).

Many industries require large quantities of materials that are taken from the Earth by the extractive industries. This may result in disruption of the geosphere from mining, dredging, and pumping of petroleum. The other major effect upon the geosphere results from the need to dispose of wastes. Scarce land may be required for waste disposal dumps and the geosphere may become contaminated with pollutants from disposal of wastes.

The biosphere is most affected by industrial activity when toxic substances are released. Other effects upon the biosphere may be indirect as the result of adverse effects upon the atmosphere, hydrosphere, or geosphere.

Industrial systems are largely dependent upon the utilization of fossil fuels, so many environmental effects are due to fossil fuel extraction and combustion. Greenhouse-warming carbon dioxide emissions, acid gas emissions, smog-forming hydrocarbons and nitrogen oxides, and deterioration of atmospheric quality from particles released from fossil fuel combustion are all atmospheric effects associated with fossil fuel combustion. Coal mining activities have the potential to release acid mine water to the hydrosphere, petroleum production can release brines or result in ocean oil spills, acid precipitation may acidify isolated lakes, and water used as cooling water in power plants may become thermally polluted. The geosphere may be disrupted by fossil fuel extraction, especially in the surface mining of coal. Coal is extracted from some areas of West Virginia by cutting off entire mountain tops overlying coal seams and dumping the overburden into valleys below in order to get to the coal. Effects upon the biosphere from fossil fuel utilization may be direct (birds coated with tar from oil spills come to mind), but are more commonly indirect, such as acidified bodies of water from acid rain resulting from sulfur dioxide emissions from coal combustion.

Agricultural activities certainly have to be considered as parts of the anthrosphere, and modern agricultural practices are part of vast agriculturally based industrial systems. Large quantities of greenhouse-warming methane are released to the atmosphere from the action of anoxic bacteria in rice paddies and in the intestines of ruminant animals. “Slash and burn” agricultural techniques practiced in some tropical countries release greenhouse gas carbon dioxide to the atmosphere and destroy the capacity of forests to sequester atmospheric carbon dioxide by photosynthesis. Enormous quantities of water are run through irrigation systems. Some of this water is evaporated and lost from the hydrosphere. The water that returns to the hydrosphere from irrigated fields picks up significant amounts of salt from the land and fertilizers applied to the land, so water salinity can become a problem. Underground aquifers become severely depleted by pumping large quantities of water for irrigation. The production of protein from livestock requires much more water overall than does the production of an equivalent amount of protein from grain. Animal wastes from huge livestock feedlots are notorious water polluters, adding oxygen-depleting biochemical oxygen demand (BOD, see Chapter 9, Section 9.3) and potentially toxic inorganic nitrogen compounds to water. The disturbance of the geosphere from crop cultivation is enormous. Raising livestock for food entails a much greater degree of land cultivation than does the cultivation of cereal grains. Agricultural production replaces entire, diverse biological ecosystems with artificial ecosystems, which causes a severe disturbance in the natural state of the biosphere. Another agricultural activity that affects the biosphere is the loss of species diversity in the raising of crops and livestock. In addition to the loss of entire species of organisms, the number of strains of organisms grown within species tends to become severely diminished in modern agricultural practice. Obviously, those varieties of crops and livestock that are mostproductive are the ones that will be used to produce grain, meat, and dairy products. However, if something

happens, such as a particular variety becoming susceptible to a newly mutated virus, alternative resistant varieties may no longer be available. Finally, the raising of transgenic crops and livestock (see Section 12.12, “Agricultural Applications of Genetically Modified Organisms”)promises profound and potentially unforeseen effects upon the biosphere.

Design of Industrial Ecosystems to Minimize Environmental Impact

From the discussion above it is obvious that industrial activity, broadly defined to include agriculture as well, has a high potential to adversely affect the atmosphere, hydrosphere, biosphere and geosphere. Inherent to the nature of industrial ecosystems, however, are measures and systems designed to minimize such impacts.

Several measures may be taken to minimize the effects of industrial ecosystems upon the geosphere. Since most of the raw materials required for manufacturing originally have to be extracted from the geosphere, the recycling of materials inherent to well designed industrial ecosystems minimizes impact upon the geosphere. The selection of materials can also be important. As an example, the mining of copper to make copper wire once widely used to carry communications signals involves digging large holes in the ground and exposing minerals that tend to release metals and acidic pollutants. The silica used in the fiber optic cables that now largely substitute for copper is simply obtained from sand. The impacts of disturbing the geosphere for food and fiber production can be minimized by some of the conservation methods and agricultural practices discussed in Chapter 11.

Well designed industrial ecosystems emit much less harmful material to the atmosphere than do conventional industrial systems. Industrial atmospheric emissions have been decreasing markedly in recent years as the result of improved technology, more stringent regulation, and requirements to release information about atmospheric emissions. One of the main classes of industrial atmospheric pollutants has consisted of the vapors of volatile organic compounds (VOCs). These have been significantly reduced by modifying the conditions under which they are used to lower emissions and by measures such as activated carbon filters to trap the vapors. The practice of industrial ecology goes beyond these kinds of measures and attempts to find substitutes, such as water-based formulations, so that volatile organic compounds need not even be used.

Years of regulation have resulted in much lowered releases of water pollutants from industrial operations. These lowered levels have been due largely to sophisticated water treatment operations that are applied to water before it is released from a plant. Desirable as these “end-of-pipe” measures are, the practice of industrial ecology goes beyond such pollution control, minimizing the use of water and preventing its pollution in the first place. One way to ensure that water pollutants are not released from an industrial operation is to completely recycle water in the system—no water out, no water pollutants.

In past years, many hazardous solid and liquid wastes have been improperly disposed to sites in the geosphere, giving rise to a large number of “hazardous waste sites,” the subject of Superfund activity in the United States. The practice of industrial ecology seeks to totally eliminate any such wastes that would require disposal. Ideally, such wastes simply represent material resources that are not properly utilized, a fact that can serve as a guideline for the prevention of such wastes.

The expenditure of energy entails the potential to cause environmental harm to the various spheres of the environment. A prime goal in the proper practice of industrial ecology is the most efficient use of the least polluting sources of energy possible. More efficient electric motors in industrial operations can significantly reduce electricity consumption. The proper design of buildings to reduce heating and cooling costs can also reduce energy consumption. Many industrial operations require heat (process heat in industrial parlance) and steam. Rather than generating these separately, they can be produced in combined power cycles along with the generation of electricity, thereby greatly increasing the overall efficiency of energy utilization.

This page titled 2.5.13.7: Environmental Impacts of Industrial Ecosystems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.7: Environmental Impacts of Industrial Ecosystems** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.8: Green Chemistry and Industrial Ecology

Green chemistry has an essential role to play in the development of successful industrial ecosystems, especially in making industrial metabolism as efficient, nonpolluting, and safe as possible. A major advantage of the practice of green chemistry to reduce environmental impact is that, ideally, it is inherently safe and clean. By using nontoxic chemicals and processes that do not threaten the environment, green chemistry avoids posing threats to the people who practice it and to the surrounding environment. Of course, these are ideals that can never be completely realized in practice, but by having these ideals as goals and making constant incremental improvements, the practice of green chemistry can become increasingly safe, environmentally friendly, and sustainable. This reduces dependence upon the command and control measures that require constant vigilance to maintain. Rather than depending upon regulations imposed from the outside to maintain its safe operation, green chemistry is much more self-regulating.

As discussed in Chapter 2 and illustrated in Figure 2.3, **Green chemistry** can be defined as the practice of chemical science and manufacturing in a manner that is sustainable, safe, and non-polluting and that consumes minimum amounts of materials and energy while producing little or no waste material. In Section 2.10, twelve principles of green chemistry were presented. Following the discussion of industrial ecology and sustainability at the beginning of this chapter, green chemistry is now covered in greater detail here. Most of the modern aspects of green chemistry are discussed in a review article on the subject.[1] The major aspects of green chemistry discussed here include the following:

1. Efficient use of matter with minimum production of wastes
2. Catalysis
3. Utilization of biological processes
4. Maximization of renewable raw materials
5. Green product design
6. Minimization or elimination of solvents, use of water where possible
7. Process intensification

Green chemistry gives prime consideration to the chemical reactions and processes by which chemicals are manufactured. One approach to making chemical synthesis greener is to use existing chemical synthesis processes but make the process itself safer and less polluting while also making the reagents required for it by greener processes. An example of the former might be to substitute a less volatile, less toxic solvent as a reaction medium for a chemical synthesis reaction. In some cases, a reagent may be made more safely by using biological processes for its preparation in place of chemical processes. A second general approach to making chemical preparations greener is to use different reagents for the synthesis that are safer and less likely to pollute.

The practice of green chemistry is largely applied to the synthesis of organic chemicals. The history of organic synthesis abounds with examples of processes that are emphatically not “green.” One example that is sometimes cited is the synthesis beginning with explosive trinitrotoluene(TNT!) of phloroglucinol, a chemical used in relatively small quantities in the fine chemicals industry. The synthesis began with oxidation by dichromate (a carcinogenic substance) in fuming sulfuric acid (a highly corrosive material that causes horrid lesions to skin) followed by reduction with iron in hydrochloric acid and heating to isolate the product. Although the quantity of product made was only about 100 tons per year, the process generated about 4,000 tons per year of solid waste containing $Cr_2(SO_4)_3$, NH_4Cl, $FeCl_2$, and $KHSO_4$. Clearly, this was not an environmentally friendly process and a major objective of green chemistry has been to find substitute pathways for synthesis such as this one.

Several key parameters are calculated in quantifying green chemistry. As discussed in Section 2.6, one of these is **atom economy** defined as the fraction of reactant material that actually ends up in final product. The higher the atom economy — ideally 100% — the greener the process. Oxidation, in organic synthesis the introduction of oxygen onto organic molecules, that uses various **oxidants** is an important process in synthesis. Oxidants are rated according to **oxygen availability**, the percentage of the mass of the oxidant molecule that is available oxygen, here represented as {O}. Theoretically, the molecule with the greatest oxygen availability is molecular O_2 rated at 100% (in practice one of the two O atoms usually ends up as water, H_2O). Hydrogen peroxide, H_2O_2, adds oxygen to an organic molecule, represented as “R”, according to the reaction,

$$H_2O_2 + R \rightarrow R(O) + H_2O$$

Since the O atom is 47% of the mass of H_2O, the oxygen availability of hydrogen peroxide is 47%. When ozone acts as an oxidant by donating one of its three O atoms,

$$O_3 + R \rightarrow R(O) + O_2$$

its oxygen availability is 33.3%, the percentage of the O_3 molecule that is one O atom.

For reduction, usually the addition of H atoms to a molecule, the concept of **hydrogen availability** may be used. When lithium hydride, molecular mass 7.94, is used as a reducing agent as in the synthesis of silane,

$$SiCl_4 + LiH \rightarrow SiH_4 + 4\,LiCl$$

all of its hydrogen is used and its hydrogen availability is 12.6% which (because of the low atomic mass of Li, hence the low formula mass of LiH) is the highest hydrogen availability of all the metal hydrides.

The Presidential Green Chemistry Challenge Awards

The U.S. Presidential Green Chemistry Challenge Awards, administered by the Environmental Protection Agency with partial sponsorship by the American Chemical Society, are made annually to recognize efforts to reduce hazards and wastes from chemical processes and to help meet pollution prevention goals. The 77 winners of these awards since 1995 are estimated to have eliminated 97 million kg of hazardous chemicals from use, prevented the release of 26 million kg of greenhouse-warming carbon dioxide to the atmosphere, and saved 80 million liters of water. The 2010 awards included the following

• The development of genetically engineered microorganisms that can convert CO_2 to higher alcohols with more than 2 C atoms that have a lower percentage of oxygen and are more hydrocarbon-like than ethanol, therefore more useful as fuels and chemical feedstocks.

• The design of a biomaterials refinery as a substitute for a petroleum refinery that uses transgenic microorganisms to convert sugars to hydrocarbon alkanes and alkenes , long-chain fatty acids, and fatty esters

• The development of a process to use hydrogen peroxide to oxidize propylene to propylene oxide, one of the most widely used organic chemical feedstocks, using an approach that reduces capital costs, energy use, and wastewater generation.

• Alterations in the synthesis of sitagliptin, the active ingredient in Merck's Januviatype 2 diabetes drug. Sitagliptin is a chiral β-amino acid and the improved approach to its synthesis uses a genetically engineered transaminase enzyme to convert a precursor ketone to the desired product with increased yield, fewer production steps,and less overall waste.

• Development of a slow-release tablet form of insecticidal spinosad that makes it useful for mosquito control in aquatic environments where biodegradation of spinosad had been a problem with earlier forms of the insecticide

---

This page titled 2.5.13.8: Green Chemistry and Industrial Ecology is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.8: Green Chemistry and Industrial Ecology** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.9: Predicting and Reducing Hazards

The reduction of hazards from chemical processes was discussed Chapter 2, Section 2.7, “Reduction of Risk: Hazard and Exposure.”The conventional approach to making chemical processes less dangerous to workers and less harmful to the environment has emphasized *exposure reduction*. In the arena of worker safety, this has involved measures such as wearing protective gear to prevent contact with hazardous chemicals. For the environment as a whole it has consisted largely of “end-of-pipe” measures to prevent release of pollutants once they are generated.

In contrast to exposure reduction, green chemistry relies upon **hazard reduction**. The first step in hazard reduction is to know what the hazards are and where they originate. Hazards may arise from the raw materials used, the media (solvents) in which chemical processes are carried out, catalysts that enable chemical reactions to occur, and byproducts. The direct hazards posed to workers in a chemical process fall into the two main categories of toxicity hazards and hazards associated with uncontrolled events such as fires and explosions.

Toxic substances are most logically classified according to their biochemical properties that lead to toxic responses. A useful means of relating toxic effects to the chemical nature of toxic substances is through **structure activity relationships**, which use computer programs to find correlations between features of chemical structure, such as groupings of functional groups, and the toxicity of the compounds. As an example, organic compounds containing the N-N=O functional group are N-nitroso compounds, a family noted for members that cause cancer. Structural features that indicate a tendency to donate methyl (-CH3) groups are also suspect because attachment of methyl groups to cellular DNA is a major mechanism in causing cancer. This hazard may be reduced by substituting alkyl groups with more carbons for the methyl group.

Three kinds of chemicals have a high priority in eliminating the toxicity hazards in green chemistry. The first such category consists of **heavy metals**, such as lead, mercury, and arsenic (a metalloid). These indestructible elements have a variety of toxic effects, such as impaired function of renal tubules in kidneys (cadmium), neurological damage (mercury), and inhibition of the production of ATP (see Chapter 7, Section 7.8, “Biochemical Processes in Metabolism”) A second major category of toxic substance that should be avoided in the practice of green chemistry consists of **lipid-solubleorganics** that are not readily degraded. These compounds often consist of relatively high molecular mass organohalides, such as PCBs, and accumulate in lipid (fat) tissue. Released into the environment, these materials can undergo **biomagnification** as they move through a food chain. **Volatile organic compounds** (**VOCs**) constitute a third class of troublesome toxic substances. These are primarily hydrocarbons, such as heptane, and low-molecular-mass organohalides, such as trichloroethylene.



Trichloroethylene

Figure6.5) known to have caused liver cancer in workers exposed to the vapor in polyvinylchloride polymer manufacture.

Chemicals that pose hazards because of their potential to undergo destructive chemical reactions fall into several often overlapping categories. **Combustible** or **flammable** substances are those that are liable to burn vigorously and destructively in air or other sources of oxygen. Hydrocarbon solvents may closely resemble gasoline in their characteristics and are highly flammable. Adding to the hazards of such materials is their volatility, which enables them to form explosive mixtures of vapor in air. Destructive petroleum refinery fires fed by hydrocarbons can occur.

Whereas combustible substances are chemical *reducers*, another category of hazardous chemical substances consists of chemical **oxidizers** that provide sources of oxygen for the reaction of reducers (see Section 13.8). One such compound is ammonium perchlorate, NH_4ClO_4, a potent oxidizer used in rocket fuel. Many chemical synthesis procedures involve oxidation steps and a variety of chemical oxidizers are used for these steps. A fourth category of hazardous chemical substances are those that are **reactive**. Explosives are prime examples of reactive substances. One of the more treacherous explosives is nitroglycerin which undergoes the following reaction when it explodes.

$$4\,C_3H_5N_3O_9 \rightarrow 12\,CO_2 + 10\,H_2O + 6\,N_2 + O_2$$

This reaction shows that the nitroglycerin molecule actually contains an excess of oxygen because elemental oxygen is released when the nitroglycerin explodes. Some structural features of molecules are known to be associated with reactivity. One example is

the close presence of oxygen and nitrogen on the same molecule, particularly where multiple nitrogen atoms are bonded together. A final category of hazardous chemical substances consists of those that are **corrosive**. In general, a corrosive substance is one that attacks materials, including even human flesh. A more specific definition is that a corrosive substance is one that produces high concentrations of either H^+ ion (strong acid) or OH^- ion (strong base).

Often hazardous chemicals belong to two or even more of the categories outlined above. An example of such a substance is concentrated nitric acid, HNO_3. In addition to its being a strong acid, this material acts as a source of oxygen, represented as $\{O\}$ in the reaction below:

$$2\,HNO_3 \rightarrow 3\,\{O\} + H_2O + 2\,NO$$

Hazardous concentrated nitric acid is used in some chemical synthesis reactions in which an acidic oxidant is needed. Among the many other potentially hazardous strong oxidants used are permanganate (MnO_4^-) and oxygen-containing compounds such as potassium dichromate, $K_2Cr_2O_7$.

The ideal way to deal with hazardous substances in the practice of green chemistry is to totally avoid making or using them. A laudable goal in principle, it is often impossible to completely avoid hazardous materials in practice. In cases where dealing with hazardous substances is necessary, every effort must be made to prevent their release, exposure to humans, or circumstances in which their hazards may be manifested.

This page titled 2.5.13.9: Predicting and Reducing Hazards is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.9: Predicting and Reducing Hazards** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.10: The E-Factor in Green Chemistry

Although atom economy, defined in Section 13.8, is a useful concept, one that is a more accurate measurement of environmental acceptability of a chemical manufacturing process is the **Efactor** defined as the following:

$$\text{E factor} = \frac{\text{Total mass of waste from process}}{\text{Total mass of product}}$$

The E factor takes into account waste byproducts, leftover reactants, solvent losses, spent catalysts and catalyst supports, and anything else that can be regarded as a waste. Its calculation depends upon what is defined as waste. For example, water is a significant byproduct of many chemical processes and is generally harmless, so its mass is usually omitted from the total mass of waste in the calculation. However, it may be included in those processes in which it is severely contaminated and difficult to reclaim in a form pure enough to use or discharge to a publicly owned wastewater treatment facility. Leftover reactant that can be easily reclaimed and recycled to the process is not included as waste whereas reactant that cannot be salvaged is counted in the waste.

The ideal E factor is 0 and higher E factors are relatively less desirable. E factors that can be tolerated depend upon the value of the product and the amount of product produced. For bulk chemicals manufactured in amounts of hundreds of thousands to millions of tons per year, tolerable E factors typically range from 1 to 5. In the fine chemical and specialty chemical industry where annual quantities are typically measured as a few thousand tons per year E factors up to around 500 may be acceptable if the value of the product is high enough to justify the cost of treating and disposal of wastes. In the pharmaceutical manufacturing industry where annual quantities generated typically are measured in tens to several hundred tons per year, acceptable E factors may be up to about 4000.

Until recently, little attention had been given to amounts of wastes produced in pharmaceutical manufacturing because the prices of the products were so high and the total amounts of wastes produced were so low. However, with the realization that even the generation of a few hundred tons per year of waste can be undesirable and costly, the pharmaceutical industry is becoming a leader in the implementation of green chemical practice. It should be noted that,although they are not considered in the calculation of E factors for pharmaceutical manufacture, annual releases of post-consumer pharmaceuticals and their metabolites are not insignificant. Of greatest concern is contamination of wastewater, some of which gets back into drinking water supplies, by pharmaceuticals and their metabolites discharged with urine or simply flushed down the drain when no longer needed. By their nature, pharmaceuticals are metabolically active and their presence in drinking water can be a concern.

The Nature of Wastes

There are wastes, and then there are wastes. Production of a few thousand tons of carbon dioxide per year may be of little concern because it can be discharged to the atmosphere, contributing to the atmosphere's burden of greenhouse gases, but negligible compared to the millions of tons released by burning fossil fuels. However, generation of a few kilograms of heavy metal wastes can be a matter of concern because of heavy metal toxicity. So it matters *whatkinds* of wastes are produced. Attempts have been made to assign an **environmental quotient**, **EQ**, to wastes where Q is a number assigned to a particular kind of waste which, multiplied times the E-factor provides in principle a means of weighting the potential harm of various kinds of wastes. Whereas E is easily measured by simple weighing, Q is a much more arbitrary number and subject to change as information is obtained regarding the potential harm of particular kinds of wastes.

This page titled 2.5.13.10: The E-Factor in Green Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.10: The E-Factor in Green Chemistry** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.11: Catalysts and Catalysis

The main components of a chemical process can be divided into the categories of catalysts, media, feedstocks, and reagents, all of which are important in the practice of green chemistry. Catalysts and catalysis are addressed in this section and the other aspects of chemical production are addressed in Sections 13.13 and 13.14.

An ideal green chemical reaction occurs with 100% atom efficiency using only reactants and no other reagents under mild conditions with only moderate input of thermal energy and without any catalysts. Unfortunately, few chemical processes meet these criteria.

In the past, especially in the synthesis of fine chemicals and pharmaceuticals where the objective has been to simply make the desired product without much consideration of waste, so-called stoichiometric reagents have been used. These have included inorganic oxidants such as MnO_2, $KMnO_4$, $K_2Cr_2O_7$; metal reductants including zinc, magnesium, sodium, and iron; and metal hydride reducing agents, especially $LiAlH_4$ and $NaBH_4$. Various organic reactions including sulfonations and nitrations employ Lewis acids (BF_3, $AlCl_3$, $ZnCl_2$) and mineral acids (HF, H_2SO_4, and H_3PO4). Many of these processes are indirect methods of adding to organic molecules hydrogen (reduction), oxygen (oxidation), carbon, and nitrogen. A problem with these reagents is the large amount of inorganic wastes produced. Where possible, it is much more desirable to employ catalysts to attach the simplest possible forms of the elements including H2 for reduction, O_2 or H_2O_2 for oxidation, CO or CO_2 for attachment of carbon, and NH_3 for attachment of N.

Instead of relying only on stoichiometric reagents, catalysts are commonly employed. As discussed in Chapter 5, Section 5.5, *catalysts*, are substances that enable reactions to proceed at significant rates without themselves being consumed. Catalysts are of great importance in the practice of green chemistry for a number of reasons including their ability to facilitate reactions and to reduce the energy required to enable reactions to proceed. There are two major approaches to contacting catalyst with a reaction mixture as shown in Figure 13.4. In **heterogeneouscatalysis** the catalyst is held on a support and the reactants flow over it. In **homogeneous catalysis** the

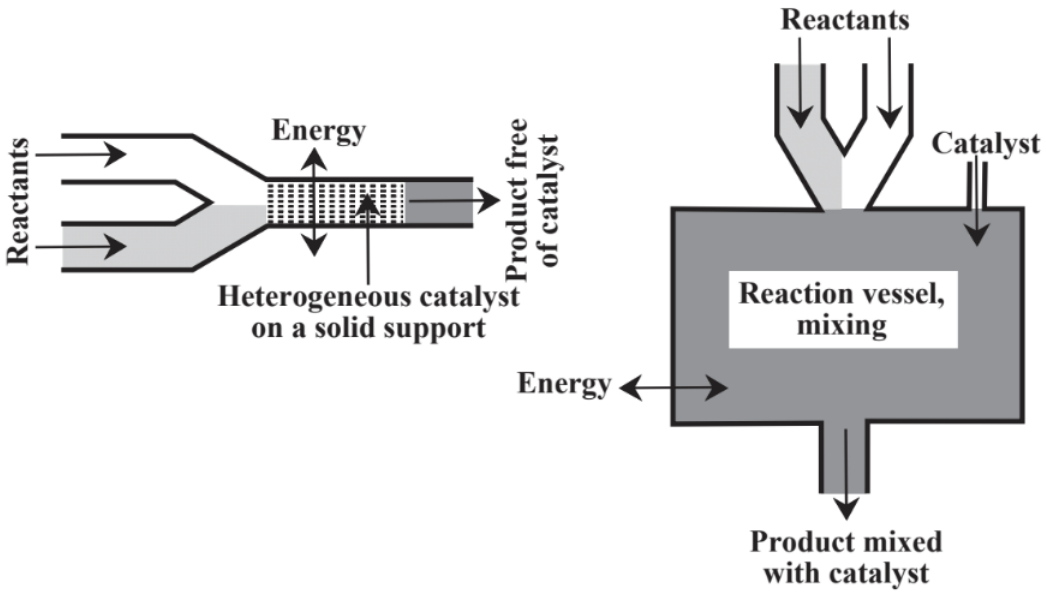


Figure 13.4. A heterogeneous catalysis (left) is held on a solid support and the reaction mixture flows over it. Inhomogeneous catalysis (right) the catalyst is mixed with the reactants and either remains in the product or is removed in a separation step. Some reactions require addition of heat energy whereas some exothermic reactions require cooling. In the practice of green chemistry energy is sometimes added as electrical current, ultraviolet radiation, microwave radiation, or ultrasound.

catalyst is placed in the reaction mixture and either remains in the product or is separated from the product in a separate step. In general, homogeneous catalysts have high activity and selectivity whereas heterogeneous catalysts are more easily recovered and recycled. Much of the activity in green chemistry has been in the area of catalysis, especially in the development of heterogeneous

catalysts that do not contaminate the product and that can be reused multiple times. Large numbers of different catalysts are used in chemical processes and their potential toxicities, production of byproducts and contaminants, recycling, and disposal are matters of considerable importance in the chemical industry.

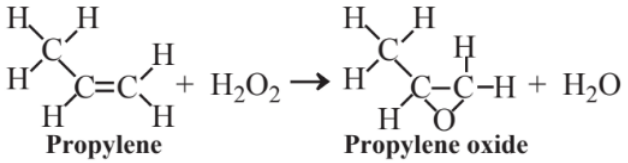
An important area of endeavor in the development of improved catalysts with respect to green chemistry is **selectivity enhancement**. Basically, this means developing a catalyst that is very selective in what it does, ideally making the right product and nothing else. A highly selective catalyst increases the percentage utilization of raw material (increased percent yield) and decreases the amount of waste byproducts from undesired side reactions.

Another important attribute of a good catalyst is related to the basic way in which a catalyst works, which is by lowering the activation energy that is required to make a reaction proceed at a significant rate. As a consequence, catalysts lower the total amount of energy that must be put into a chemical process to get it to occur. Lowered energy requirements are a basic part of the practice of green chemistry and in this respect highly efficient catalysts can be extremely beneficial in reducing energy consumption and in so doing lowering costs and environmental impact.

Numerous green industrial chemical reactions have been developed in recent years, most employing some sort of catalyst. An example is the synthesis of the widely used industrial chemical propylene oxide starting with elemental H_2 and O_2. Around 1985 it was discovered that propylene could be oxidized to propylene oxide with 30% hydrogen peroxide using a titanium silicate catalyst, but the process was not economic because of the high cost of hydrogen peroxide. A Green Chemistry Presidential Challenge prize was awarded in 2007 for an economical process to make hydrogen peroxide by combining H_2 and O_2 in a gas mixture at levels of H_2 directly below the lower flammability limit of H_2 using a catalyst composed of palladium-platinum nanoparticles:

$$H_2 + O_2 \rightarrow 2\,H_2O_2$$

As a second step in the process the hydrogen peroxide is reacted with propylene to produce propylene oxide with water as the only byproduct:



(13.11.2)

Figure 13.5).

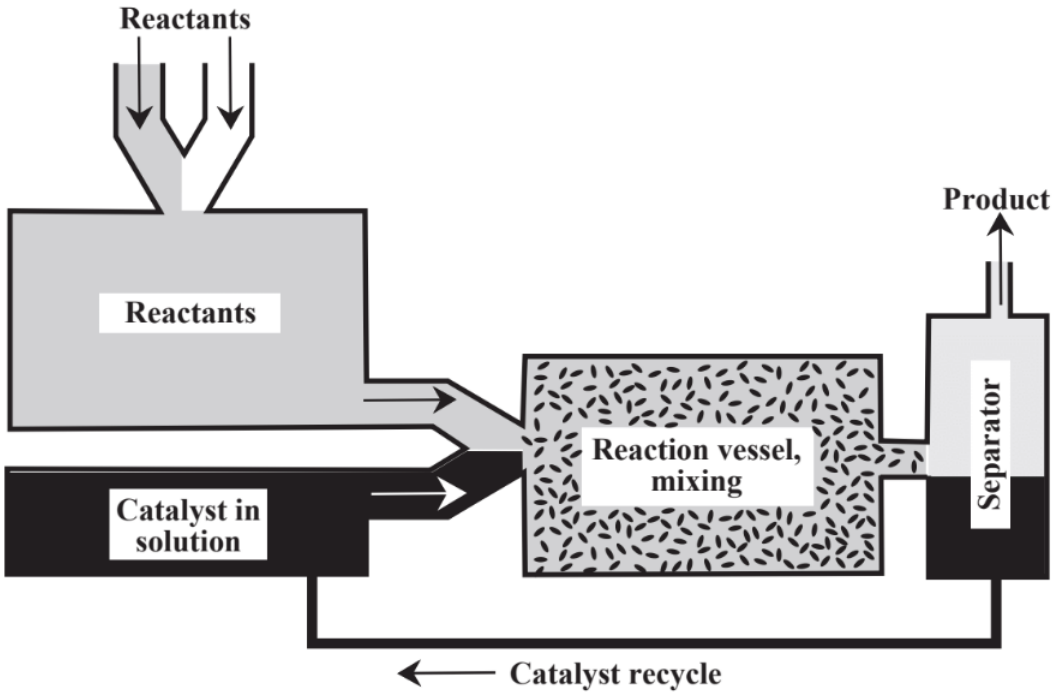


Figure 13.5. Illustration of a biphasic catalyst system. Reactants in one solution contact catalyst in another immiscible solution. After the

reaction has had a chance to go to completion, the two immiscible phases are separated and the catalyst solution recycled. In the reaction vessel this kind of system provides the superior qualities of a homogeneous catalyst but with the product/catalyst separability of a heterogeneous catalyst.

This page titled 2.5.13.11: Catalysts and Catalysis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.11: Catalysts and Catalysis** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

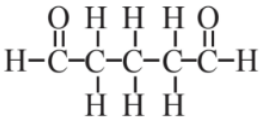
2.5.13.12: Biocatalysis with Enzymes

Nature has provided some ideal catalysts in the form of enzymes, the use of which offers some substantial advantages in green chemistry. The most obvious advantage is that enzymes have evolved to work under the mild conditions in which organisms function, especially moderate temperatures and physiological pH. Unlike precious metal catalysts that are commonly used, enzymes are made renewably from biomaterials. Enzymes have generally high activities and are highly selective in the chemical processes that they carry out. Whereas conventional organic syntheses often require activation or protection of functional groups, both of which may consume reagents and hence produce more wastes, these measures are often not needed with enzymes. The result is that biocatalyzed reactions can frequently be carried out with relatively fewer steps and less wastes making them more attractive environmentally and economically.

Although some biocatalyzed reactions have been used for production of chemicals for centuries — the production of ethanol by yeast fermentation of sugar comes immediately to mind— relatively recent advances in biotechnology have greatly increased their versatility and utility. One such advance has been with recombinant DNA in which enzymes that perform specific functions may be transferred between organisms. The other major advance has been with **directed evolution** in which the amino acid sequences in enzymes produced by genes are shuffled randomly and large numbers of the products are sampled for their activity, particularly for carrying out a particular biochemical synthetic step. This may be done within living cells, but can be carried out on a much larger scale outside of cells, a process called ***invitro* evolution**. Obviously, most of the enzymes produced by this technique are not superior, or are even useless, but out of the enormous numbers generated, some will be superior. *In vitro* evolution is being carried out to provide enzymes with properties such as improved or different catalytic activity, catalytic specificity, thermostability, and pH optima that can be used in industrial, medical, or agricultural applications. It is proving particularly useful in developing enzymes that act upon compounds that do not occur in nature and for which enzymes have not evolved through natural evolution.

Immobilized Enzyme Catalysts

Disadvantages of catalytic enzymes include often low stability, limited storage times, difficult recovery, and product contamination, almost always a consideration with homogeneous catalysts. The stability and recyclability of enzymes may often be enhanced by immobilizing them on solid supports. A common means of immobilizing enzymes is by their precipitation from the fermentation broth in which they are generated by a buffer such as ammonium sulfate and stabilizing the precipitated aggregates with a reagent possessing two bonding functional groups, commonly glutaraldehyde a reagent that is widely used for cross-linking proteins:



Glutaraldehyde

The precipitation and cross-linking of the enzyme combines its isolation and immobilization into a single step. Enzymes prepared in this way are usually highly productive, stable and resistant to denaturation (loss of enzyme function by structural alteration) due to exposure to organic solvents, heating, or breakdown to shorter peptide chains or amino acids by the action of proteolytic enzymes (proteolysis).

Reduction in Synthesis Steps with Enzyme Catalysts

Typical synthesis of an organic compound, especially one as complicated as many pharmaceutical agents, may involve multiple steps. Because of factors such as product loss, the need to protect and deprotect functional groups, and generation of wastes from each step, these multistep syntheses tend to build high E factors overall. The ideal synthesis is a “one-pot” process in which all steps are carried out in the same operation without the need to isolate intermediates. Living cells are often “one-pot” synthetic factories, so it is natural to look to enzymes to accomplish the same thing in the laboratory and in chemical production using multiple enzymes ina single container and for a single multiple-step operation. Accomplishment of such a process can be complicated by incompatibility of different enzymes and the different conditions under which enzymes operate although in general they operate in water under ambient conditions compatible with life. (The relatively recent discovery of organisms that live under

hot, extreme conditions on the deep ocean floor raises some interesting possibilities for the isolation of enzymes that might function under unusual conditions, particularly elevated temperatures.)

Enzyme Catalysts and Chirality

As shown by the example of the herbicide mecoprop in Figure 13.6, **chiral molecules** are three-dimensional molecules with structures such that a molecule cannot be directly imposed on its mirror image. Chiral molecules have different groups arranged around an atom, usually of carbon, that constitutes a chiral center. Two chiral molecules of the same compound are called **enantiomers** commonly designated as R and S. The physical and chemical properties of enantiomers are generally identical, having the same melting points, boiling points and solubilities. However, one enantiomer of a compound may fit exactly with an enzyme active site whereas the other does not. This results in often markedly different biochemical properties of enantiomers and consequently completely different environmental and toxicological behaviors. One enantiomer of a chiral pharmaceutical may function very well whereas the other does not work at all or may even be toxic. Biochemical differences between enantiomers may be especially pronounced for pesticides. For example, the R enantiomer of herbicidal mecoprop (Figure 13.6) kills weeds very effectively whereas the S enantiomer is inactive; therefore, the pure R enantiomer is now marketed as a herbicide in place of the racemic mixture with the S enantiomer.

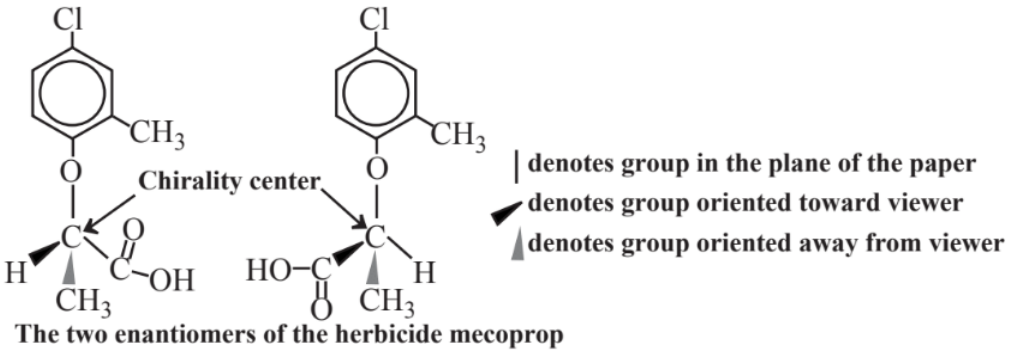


Figure 13.6. Illustration of the chiral herbicide mecoprop. Only the R enantiomer has herbicidal activity and is now marketed in the enantiometric pure form

Normally when a chiral compound is synthesized by conventional chemical means a racemic mixture of the two enantiomers is produced. Because of their essentially identical chemical properties they are very hard to separate. However, it is possible to produce enantiomerically pure chemical compounds with appropriate enzymatic catalysts. Furthermore, it is possible to use enzymatic catalysts to convert racemic mixtures of compounds to enantiomerically pure forms. In one of the larger such industrial operations, BASF now uses enzymatically catalyzed processes to prepare enantiomerically pure amines in thousand-ton quantities.

Nature's catalysts, the enzymes in organisms, are experts in carrying out chemical processes efficiently under mild conditions. In consideration of this fact, a great deal of attention is being devoted to using organisms, especially bacteria, to carry out chemical processes. By splicing desired genes for making specific enzymes into bacteria so that they will carry out desired reactions, genetic engineering is making an increasing contribution to the development of enzyme-catalyzed green chemical processes.

Chemists are trying to use enzymes as models for synthetic catalysts that have performance characteristics of enzymes, but which are much simpler and work under conditions that would destroy enzymatic catalysts. A promising area in which this might be accomplished is the use of iron-containing catalysts to oxidize alkene (C=C) groups in organic compounds using relatively mild hydrogen peroxide reagent, H_2O_2. Organisms accomplish this task using catalysts in which the Fe^{2+} ion is bonded by four N atoms in relatively large heme porphyrin molecules. The same oxidation has now been accomplished with a catalyst in which Fe^{2+} is bound by four N atoms by an organic molecule with the formidable name of N,N'-dimethyl-N,N'bis(2-pyridylmethyl)-ethylenediamine as shown in Figure 13.7. A big advantage of this catalyst that is shared with enzyme catalysts that enable peroxide oxidations is that it does not cause the decomposition of hydrogen peroxide as do a number of synthetic catalysts.

This page titled 2.5.13.12: Biocatalysis with Enzymes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.12: Biocatalysis with Enzymes** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.13: Energizing Chemical Reactions and Process Intensification

One of the most important aspects of green chemistry is the enhancement of the speed and degree of completion of chemical reactions. One of the ways that this is done is by lowering the activation energy required to enable a reaction to proceed. That is what catalysts do as discussed in the preceding section. The other way to enhance a reaction is by adding energy as discussed in this section.

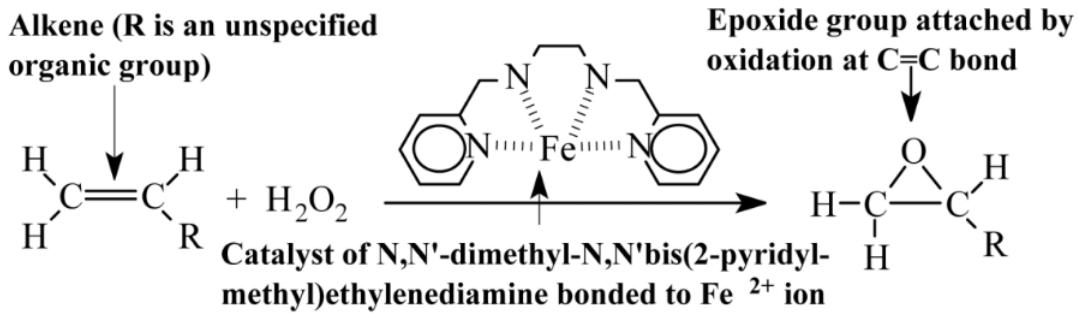


Figure 13.7. Action of an iron-containing compound modeled after iron-based enzymes as a catalyst to bring about the oxidation of an alkene hydrocarbon group with hydrogen peroxide

The most straightforward means to add energy to a reaction is by heating the reaction mixture. On an industrial scale this is commonly accomplished with coils of tubing immersed in the reaction mixture that are heated with steam passing through the coils. Heating by passing a current of electricity through electrically-resistant coils is also a means of adding energy to a chemical system. Much of the effort in green chemistry has been devoted to finding more sophisticated ways of energizing chemical systems.

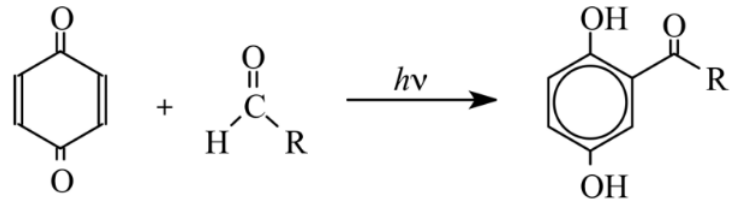
Microwaves can be used to add energy to reactions to enhance reaction rates. Microwaves are electromagnetic radiation with wavelengths of 1 cm to 1 m (frequency 30 GHz to 300 Hz). To avoid interference with microwave bands used in communication, industrial and household microwave generators commonly operate at 2.45 GHz. Microwaves are absorbed by polar molecules, such as those of water, causing rapidly repeating re-orientation of the molecules in a microwave field. The result is a high input of energy directly into substances subjected to microwaves thereby adding energy and speeding up reactions. Microwave energy can be put directly into relatively small volumes of reaction media, reducing material requirements and minimizing wastes. Microwaves can be used to enhance reactions in (1) water media, (2) polar organic solvents such as dimethylformamide, and (3) media-free reactions, such as mixed solid reactants.

Sonochemistry adds energy by subjecting a reaction medium to ultrasound energy at frequencies between 20 and 100 KHz which introduces very high energy pulses into the medium. Commonly, the ultrasound is produced by the piezoelectric effect through which crystals of substances such as ceramic-impregnated barium titanate are subjected to rapidly reversing electrical fields converting the electrical energy to sound energy with an efficiency that can reach 95%. An advantage of sonochemistry is that it can introduce high energy into microscopic regions enabling reactions to occur without appreciably heating the reaction medium.

Electrochemistry by the passage of a direct current of electricity through a reaction medium can cause both reductions and oxidations to occur. Reduction, the addition of electrons, *e-*, occurs at the relatively negatively charged cathode, and oxidation, the loss of electrons, at the relatively positively charged anode. Electrochemical oxidation and reduction can be controlled by the electrical potentials applied, by the media in which they occur, and by the electrodes used. Because the addition of electrons to the reaction medium (reduction) and their accompanying removal (oxidation) does not add matter, electrolytic syntheses meet the goals of green chemistry. The electrolytic production of oxygen and of hydrogen, a non-polluting fuel and valuable raw material, is shown in Section 13.15 and Figure 13.10.

Photochemical reactions use the energy of photons of light or ultraviolet radiation to cause reactions to occur. The energy, E, of a photon of electromagnetic radiation of frequency,ν, is $E = h\nu$, where *h* is Planck's constant. Since a photon can be absorbed directly by a molecule or a functional group on a molecule, the application of electromagnetic radiation of the appropriate energy to a reaction medium can introduce a high amount of energy into a reactant species without significantly heating the medium. Photochemical energy can be used to cause synthesis reactions to occur more efficiently and with less production of waste

byproducts than nonphotochemical processes. One example is the acylation of benzoquinone with an aldehyde to produce an acyl hydroquinone, an intermediate used to make some specialty polymers:



Benzoquinone **An acylhydroquinone (R is a hydrocarbon group)** (13.13.1)

This reaction occurs with 100% atom economy. Unlike the standard Friedel-Crafts type of reaction, which utilizes the catalytic effect of Lewis acid-type acidic halides, particularly aluminum chloride, $AlCl_3$, the photochemical process does not require catalytic substances which tend to be reactive and moisture- and air-sensitive.

A reaction participant does not have to absorb a photon directly to undergo a photochemically induced reaction. In some cases photochemically reactive species are added to the reaction mixture to absorb photons, then produce reactive excited species or free radicals that carry out additional reactions. An example of this is provided with hydrogen peroxide, which absorbs photons.

$$H_2O_2 + h\nu \rightarrow HO \cdot + HO \cdot$$

to produce reactive hydroxyl radicals that react with a number of other species.

Process Intensification and Increased Safety with Smaller Size

Process intensification can be employed with continuous-flow reactors (Figure 13.8) used to intensify chemical processes and enable increased output of product with a smaller footprint of apparatus. This is especially the case when continuous flow is combined with heterogeneous catalysis and energy input. A big benefit to such reactors from the green chemistry viewpoint is increased safety. If something goes wrong in a large batch reactor, in the worst case an accident such as an explosion or fire with a large amount of material may occur. With a continuous-flow reactor the problem can be confined to the small volume of the reactor and the process can be shutdown immediately by stopping the inflow of the reactants.

This page titled 2.5.13.13: Energizing Chemical Reactions and Process Intensification is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.13: Energizing Chemical Reactions and Process Intensification** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.14: Solvents and Alternate Reaction Media

Chemical reactions are often carried out in **media**, usually organic **solvents** or water, to provide a medium in which feedstocks and reagents can dissolve and come into close, rapid contact at the molecular level. A good solvent for chemical synthesis is one that enables facile product separation and is amenable to purification and reuse with minimum loss. Substances dissolved in a solvent are **solvated** by binding of the solvent to the molecules or ions of the dissolved substance, the **solute**. Solvation of reactants and products often plays an important role in determining the kinds and rates of reactions. Many organic feedstocks and reagents are not soluble in water or are decomposed by it, so organic solvents including hydrocarbons, chlorinated hydrocarbons, and ethers have to be used as reaction media.

Organic solvents cause several problems in chemical synthesis. Particularly because of problems associated with their containment, recovery, and reuse, organic solvents especially are major contributors to undesirably high E factors. Many of the environmental and health problems associated with making chemicals are the result of the use of organic solvents as media. Hydrocarbon solvents will burn and hydrocarbon vapors in air are explosive. Although many hydrocarbon solvents are not particularly toxic, some can cause the condition of peripheral neuropathy (damage to peripheral nerves such as those in feet and legs), and benzene is regarded as a carcinogen thought to cause leukemia. Released to the atmosphere, hydrocarbons can also participate in photochemical processes leading to the formation of photochemical smog (see Chapter 10, Section 10.11).

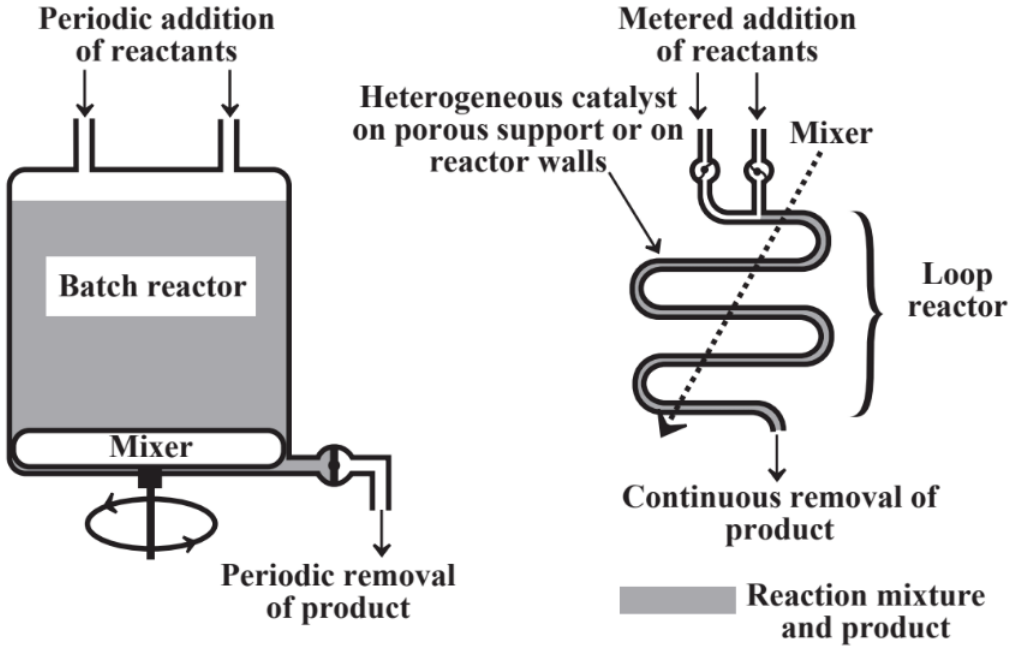
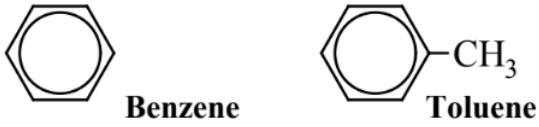


Figure 13.8. A batch reactor (left) involves mixing a large amount of reactants, often with a homogeneous catalyst, and allowing the reaction to take place for the required amount of time with the result that if anything goes wrong a large amount of material is spoiled and a considerable hazard may develop. In a continuous-flow reactor (right) only a small amount of material is reacting at any given time, heterogeneous catalysts can be used, energy input can be accurately regulated, and the process can be stopped at any time if something goes wrong.

One approach to making chemical synthesis processes greener is to replace specific solvents with less hazardous ones. For this reason, toxic benzene solvent is replaced by toluene wherever possible. As shown by their structural formulas below, toluene has a methyl group, $-CH_3$, that benzene does not possess. The methyl group in toluene can be acted upon by human metabolic systems to produce a harmless metabolite (hippuric acid) that is eliminated in the urine, whereas metabolic processes acting upon benzene convert it to a toxic intermediate that can react with cellular DNA.



Benzene Toluene

As another example of solvent replacement, *n*-hexane, which can cause peripheral neuropathy in exposed individuals, can be replaced with 2,5-dimethylhexane, which does not cause this condition, for reactions where the higher boiling temperature of the latter compound is not a problem.

There is significant interest in reaction media other than organic solvents. The ultimate approach to eliminating problems with solvents in chemical synthesis is to do reactions without solvents of any kind. Some reactions can be performed in which the reactants are simply mixed together or are held on solid supports, such as clays. Microwave heating of such reaction mixtures has proven effective in providing energy to enable reactions to occur rapidly. However, in many cases that is not possible and solvents are required. Some alternative solvents are discussed below.

Water Solvent

Although many reagents are reactive with water making its use impossible, where applicable, the greenest solvent for green chemical processes is water. Water is abundant, cheap, not toxic, and does not burn. Because of its polar nature and the ability to form hydrogen bonds (see Chapter 9, Section 9.1 and Figure 9.1), water is an especially good solvent for ionic compounds — acids, bases, and salts. Water is particularly useful as one of the solvents in biphasic catalysis described in Section 13.11. Normally the catalyst is held in the water phase and the product in a water-immiscible organic solvent, which allows facile separation of the catalyst after the reaction is complete.

Because of water's many advantages, significant efforts have been made in replacing organic solvents used for reaction media with water. Although water does not appreciably dissolve many nonpolar organic compounds, in some cases these materials may be suspended as very small colloidal particles in water, enabling close enough contact of organic materials to undergo reactions. Water is a good solvent for some of the biological materials, such as glucose, now favored as chemical feedstocks where they can be used.

Carbon Dioxide Solvent

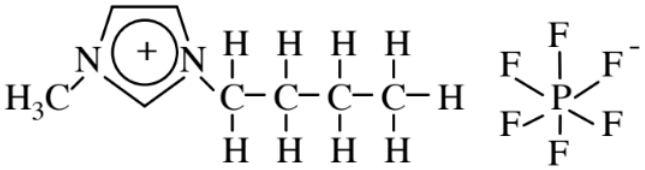
At a high pressure above 73.8 atm (73.8 times normal atmospheric pressure at sea level) and a temperature exceeding 31.1° C, carbon dioxide becomes a **supercritical fluid**, a relatively dense state of matter in which there is no longer a distinction between liquid and gas. A good solvent for organic compounds, supercritical carbon dioxide can be used as a reaction medium for organic chemical reactions. An advantage of supercritical carbon dioxide in this application is that its viscosity is only about 1/30 that of common liquid organic solvents, which enables reactant species to move much faster through the fluid, thus speeding the reactions that they undergo. At temperatures and pressures somewhat below those at which carbon dioxide becomes critical, it exists as separate gas and liquid phases while retaining many of the solvent properties of supercritical carbon dioxide. Under these conditions carbon dioxide is called a **dense phase fluid**, a term that also encompasses supercritical fluids.

Adjustment of the composition and conditions under which dense phase fluid carbon dioxide is maintained can provide significant variations in its solvent properties and adjustment of its ability to act as a reaction medium. In addition to variations in temperature and pressure, dense phase fluid carbon dioxide may be mixed with small quantities of other solvents (modifiers), such as methanol, to further vary its solvent properties.

In addition to its solvent properties, dense phase fluid carbon dioxide offers the advantage of low toxicity and low potential for environmental harm (the small amounts of greenhouse gas carbon dioxide released from its application as a solvent are negligible compared to quantities released from combustion of fossil fuels). A big advantage of dense phase fluid carbon dioxide is its volatility, meaning that it separates readily from reaction products when pressure is released. Furthermore, carbon dioxide released from a reaction mixture can be captured and recycled for the same application. Carbon dioxide can be obtained at low cost from biological fermentation processes.

Ionic Liquid Solvents

Ionicliquids present another alternative to organic solvents for use as media for chemical synthesis. Inorganic salts consisting of ions, such as NaCl composed of Na^+ and Cl^-ions, are normally hard, high-boiling solids. However, when one or both of the ions are composed of large charged organic molecules, as shown by the cation in the example of an ionic liquid below,



1-Butyl-3-methylimidazolium hexafluorophosphate

the salts can be liquids at room temperature and are called ionic liquids. These materials have the potential to act as suitable media in which substances can be dissolved and undergo reactions, and active research is underway to explore this possibility. There is an enormous variety of such ionic liquids with widely varying solvent properties because of the large number of kinds of ions that can be combined leading to almost limitless possibilities for various ionic liquids. An interesting possibility that has been tried experimentally is to use a mixed ionic liquid/supercritical carbon dioxide single fluid phase in which the reaction proceeds with a homogeneous catalyst followed by reduction of pressure that causes the supercritical carbon dioxide and ionic liquid to separate into two phases with the catalyst remaining in the ionic liquid — hence available for reuse — and the product in the supercritical carbon dioxide.

This page titled 2.5.13.14: Solvents and Alternate Reaction Media is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.14: Solvents and Alternate Reaction Media** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.15: Feedstocks and Reagents

Feedstocks

Feedstocks are the main ingredients that go into the production of chemical products. As discussed below, feedstocks may be acted upon by reagents, and often there is some overlap between the two categories of materials. Feedstocks are so important in the practice of green chemistry that much of Chapter 14, “Feeding the Anthrosphere: Utilizing Renewable and Biological Materials,” is devoted to renewable feedstocks. Here they are introduced briefly as they relate to green chemistry and sustainability.

There are three major components of the process by which raw materials from a source are obtained in a form that can be utilized in a chemical synthesis, then converted to a product. The first of these is the **source** of the feedstock, an aspect that has a number of environmental and sustainability considerations. This may consist of a depleting resource, such as petroleum, in which case the lifetime of the resource and the environmental implications of obtaining it must be considered. From the standpoint of sustainability it is preferable to use recycled materials as feedstocks, although the availability of recycled materials suitable for this purpose is limited. A third source that is very desirable consists of renewable resources, particularly from materials made by photosynthesis and biological processes.

The second major aspect of converting feedstocks to final products is separation and isolation of the desired substance. An example of this step is the isolation of specific organic compounds from crude oil to provide a feedstock for organic chemical synthesis. It may be necessary to process raw materials from a source to convert it to the specific material used as a feedstock for a chemical process. Often most of the environmental harm in providing feedstocks comes during the isolation process, in large part because of the relatively large amount of waste material that often is generated in obtaining the needed feedstock.

The world chemical industry has been built primarily on fossil carbon feedstocks. Much of the impetus for the organic chemical industry was built during the late 1800s and early 1900s on the basis of organic chemicals isolated from the coal tar byproduct of coal coking. Later petroleum and natural gas became the basis of the petrochemicals industry which has produced enormous quantities of polymers, plastics, synthetic rubber, and thousands of other kinds of chemicals. Eventually this reliance on depleting fossil carbon resources must end. Therefore, one of the main goals of green chemistry has become a shift toward renewable feedstocks — biomass produced by photosynthesis — and to renewable reaction media, specifically water and supercritical and pressurized liquid carbon dioxide. In addition to being renewable, such feedstocks offer advantages over petroleum in that they are not toxic and they, and most of their products and intermediates, are biodegradable.

Biomass Feedstocks

The ongoing shift to a biobased economy offers the opportunity to design their processing and products on the basis of green chemistry and sustainability. Rather than petroleum refineries, biorefineries are used to process biobased feedstocks. Initially, the feedstocks used have been from grains of corn and oil seeds, such as soybeans. Cornstarch is hydrolyzed to sugars that are converted to alcohols by fermentation processes and to biodiesel fuel by esterification of the fatty acids in the oil seed lipids. Some commodity chemicals such as lactic acid are also made from grain sources. However, these are inefficient ways to utilize biomass, they compete with the consumption of grain for food, and the maximum production of grain requires intensive cultivation (often of marginal land), heavy use of fertilizer, and strong reliance on pesticides. A much more renewable approach is to utilize lignocellulose, the structural material of plants (such as wheat straw) either from dedicated crops or preferably from crop byproduct biomass or highly productive algae growing in water. Although attempts are being made to break down these lignocellulose materials with enzymes to produce fermentable sugars, a preferable pathway is the thermochemical route. This can be by pyrolysis of the biomass to produce a variety of organic liquids, some gas and a carbon residue that can be burned to provide heat for the pyrolysis. Alternatively, biomass can be reacted directly with hydrogen to produce liquid and gas products. Both pyrolysis and hydrogenation produce a large variety of liquids including oxygenated organics that can be run through a biorefinery. As discussed in Chapter 15 on energy, the best approach is thermochemical gasification of the biomass to produce a synthesis gas mixture of CO and H_2 that can be used with well known technology to synthesize gasoline, diesel fuel, aircraft fuel, alcohols, substitute natural gas, and various organic chemical feedstocks.

Genetic engineering can be very useful in producing biomass for feedstocks. One area in which there is much room for improvement is in enhanced efficiency of photosynthesis. Crops can be bred to increase the amount of byproduct biomass along

with the grain they produce. Dedicated crops can be developed for the production of large quantities of biomass, alone. This has already been done using conventional plant breeding techniques to develop rapidly growing hybrid poplar trees that produce large quantities of lignocellulosic wood.

Carbohydrate Feedstocks

The most abundant biomass feedstocks are carbohydrates. It follows that one of the most promising pathways to obtaining useful raw materials and fuels from biomass is their synthesis directly from carbohydrates. As an example of a feedstock chemical that can be synthesized from fructose, a monosaccharide made in abundance from cornstarch and sucrose (common table sugar) is dimethylfuran (Figure 13.9) an oxygen-containing cyclic organic compound that has most of the desirable properties of hydrocarbons as a fuel and raw material. As shown in Figure 13.9, fructose and dimethylfuran have similar structural formulas and the conversion is largely a matter of removing oxygen from fructose by reaction with hydrogen. This synthesis and monosaccharide glucose as a feedstock are discussed in Chapter 14.

Green chemistry and the inevitable shift away from petroleum hydrocarbon feedstocks to biobased materials will cause a massive realignment of the chemical industry. Rather than manufacturing nondegradable materials using environmentally unfriendly technologies with depletable feedstocks, the new paradigm will employ green catalytic processes to convert biomass feedstocks to biocompatible (non-toxic, biodegradable) products.

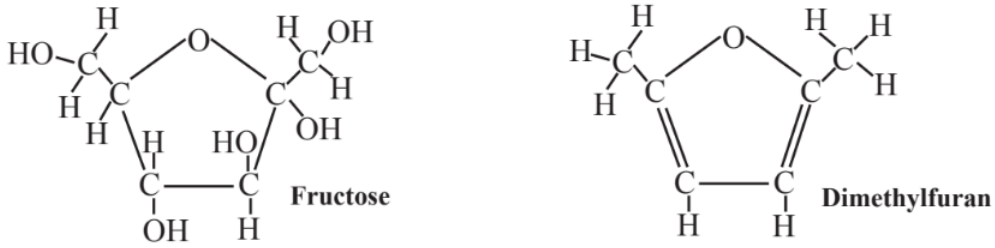


Figure 13.9. Fructose, a carbohydrate produced in abundance by plants, can be used to make dimethylfuran, a compound that resembles hydrocarbons in many of its properties and that can be used as a motor fuel and as a feedstock for organic synthesis.

Reagents

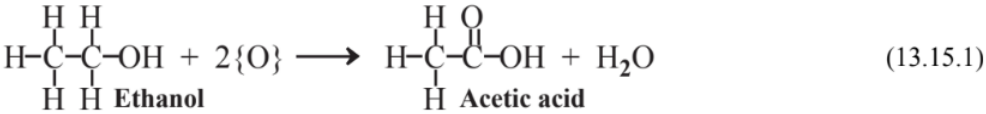
The term **reagents** is used here to describe the substances that act upon basic chemical feedstocks to convert them to new chemicals in synthetic processes. The kinds of reagents used have a very strong effect upon the acceptability of a chemical process with respect to green chemical aspects. Much of the work that has been done in developing and using green reagents has involved organic chemical processes, many of which are beyond the scope of this book. However, some of the general aspects of chemical reagents from a green chemical perspective are discussed here.

The most obvious characteristic required of a good chemical reagent is that it do what it is supposed to do completely and at an acceptable rate. A reagent with a high **productselectivity** produces a high percentage of the desired product with a low percentage of undesired byproducts. Another desirable characteristic of a good reagent is high **productyield** meaning that most of the feedstocks are converted to product. The use of reagents that provide high selectivity and yield means that less unreacted feedstock and byproduct material have to be handled, recycled, or disposed.

One of the most common measures taken in implementing green chemical processes is selection of alternative reagents. The criteria used in selecting a reagent include whether or not itis available, how efficient it is, and the effect that it has. Important considerations with the chemical transformation are whether it is stoichiometric or catalytic, the degree to which it is atom economical, and the quantities and characteristics of any wastes produced.

Reagents for Oxidation and Reduction

One of the main kinds of reactions for which reagents are used is **oxidation**, which usually consists of the addition of oxygen to a chemical compound or a functional group on a compound. (See Chapter 5, Section 5.7 for a discussion of oxidation and its accompanying phenomenon, reduction and section 13.8 regarding oxygen availability in oxidants and hydrogen availability in reductants.) An example of an oxidation reaction is the conversion of ethanol to acetic acid,



(13.15.1)

where {O} is used to represent oxygen from some unspecified oxidant. Oxidation is one of the most common steps in chemical synthesis. A number of reagents are used as oxidants. Some of these reagents, such as potassium dichromate, $K_2Cr_2O_7$ are dangerous (dichromate salts are considered to be carcinogenic when inhaled for prolonged periods of time) and leave troublesome residues that require disposal.

Because of problems with oxidants that are commonly used, a major objective in the practice of green chemistry is to use more benign oxidants. Alternatives to the more traditional oxidant reagents include molecular oxygen (O2), ozone (O3), and hydrogen peroxide (H_2O_2), usually used with a suitable catalyst that enables the oxidation reaction to occur. Under the right conditions, hydrogen peroxide can be used as an alternative to elemental chlorine, Cl_2, a strong oxidant used to bleach colored materials, such as paper pulp and cloth. Since chlorine is toxic (it was used as a poison gas in World War I) and has a tendency to react with organic compounds to produce undesirable chlorinated organic compounds, hydrogen peroxide is a much preferable bleaching agent.

In contrast to the usually harsh conditions under which chemical oxidations are carried out, organisms carry out biochemical oxidations under mild conditions. In so doing, they use monooxygenase and peroxidase enzymes that catalyze the oxidizing action of molecular oxygen or hydrogen peroxide. An area of significant interest in green chemistry is to perform such oxidations in biological systems or to attempt the use of catalysts that mimic the action of enzymes in catalyzing oxidations with molecular oxygen or hydrogen peroxide.

Reduction, which consists of loss of O, gain of H, or gain of electrons by a chemical species is also a common operation in chemical synthesis. As is the case with oxidants, the reagents used to accomplish reduction can pose hazards and produce undesirable byproducts. Such reductants include lithium aluminum hydride ($LiAlH_4$) and tributyl tin hydride.

Electrons as Reagents for Oxidation and Reduction

As an alternative to the potentially troublesome oxidation and reduction procedures using reagents, **electrochemistry** provides a reagent-free means of doing oxidation and reduction. This is possible because an electrical current consists of moving electrons and oxidation consists of electron removal from a chemical species and reduction is addition of an electron. The passage of an electrical current between metal or carbon graphite electrodes through a solution resulting in oxidation and reduction reactions is called **electrolysis**. Consider the simplest possible case of electrolysis, that of water containing a non-reactive salt, such as Na_2SO_4, shown in Figure 13.10.

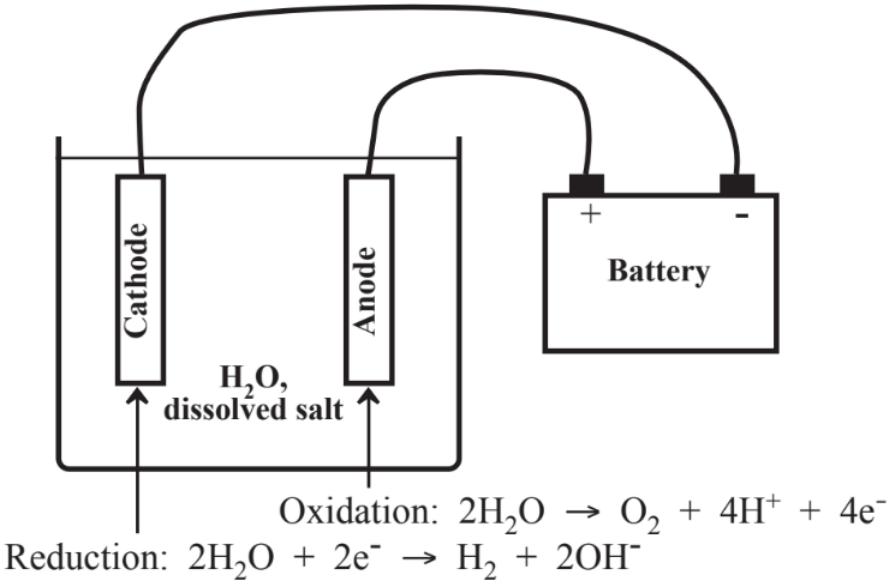


Figure 13.10. Apparatus for electrolysis in which a direct current of electricity is passed through a reaction medium, in this case water with a dissolved salt to make it electrically conducting. Reduction occurs when electrons are added to the medium at the cathode and oxidation when they are removed at the anode. Electrolysis is a reagent-free way of doing oxidation and reduction.

At the **cathode**, where electrons (e-) are pumped into the system and where reduction occurs, reduction of water occurs releasing H_2,

$$2\,H_2O + 2\,e^- \rightarrow H_2 + 2\,OH^-$$

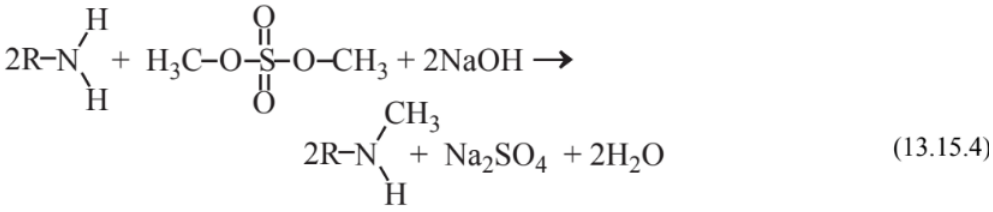
and at the **anode** where electrons are removed, O_2 is released as the water is oxidized:

$$2\,H_2O \rightarrow O_2 + 4\,H^+ + 4\,e^-$$

In the setup shown, H+ion generated at the anode OH-ion generated at the anode diffuse through the solution and react upon contact to produce water again. At the cathode, a dissolved chemical species could be reduced directly or the hydrogen generated could add to a species, reducing it. And at the anode another species could be oxidized directly by loss of electrons or the oxygen generated could add to a species, oxidizing it.

Miscellaneous Reactions with Reagents

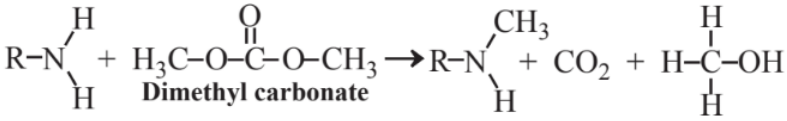
Reactions other than oxidation and reduction are carried out by reagents. As an example of a commonly performed reaction that normally requires potentially troublesome reagents, consider **alkylation** with alkylating reagents in which an alkyl group, most frequently the -CH_3 (methyl) group, is added to an atom on an organic compound. The methylation reaction,



(13.15.4)

shows attachment of a methyl group to an amine group, -NH_2, that is part of an unspecified molecule represented “R.” Methylation of nitrogen is used in a number of chemical syntheses including preparation of analgesics such as Ibuprofen. The dimethyl sulfate reagent used to accomplish the methylation poses toxicity problems in that it is a suspect human carcinogen. The reaction also produces a byproduct of Na_2SO_4, which if contaminated with dimethyl sulfate reagent may pose disposal problems.

Dimethyl carbonate prepared by reacting methanol, CH_3OH, with carbon monoxide, CO, in the presence of elemental oxygen and a copper salt catalyst has been developed as a green alternative to dimethyl sulfate as a methylating reagent. When dimethyl carbonate acts as a methylating agent,



(13.15.5)

methanol and innocuous carbon dioxide are generated as byproducts. The methanol can be recirculated through the system to generate additional dimethylcarbonate reagent.

This page titled 2.5.13.15: Feedstocks and Reagents is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **13.15: Feedstocks and Reagents** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.13.16: Literature Cited and Supplementary References

1. Sheldon, Roger A., “E Factors, Green Chemistry and Catalysis: An Odyssey,”*Chemical Communications*, 3352-3365 (2008).

Ahluwalia, V. K., M. Kidwai, *New Trends in Green Chemistry*, Kluwer Academic, Boston, 2004.

Allen, David T., and David R. Shonnard, *Green Engineering: Environmentally Conscious Design of Chemical Processes*, Prentice Hall, Upper Saddle River, NH, 2002.

Anastas, Paul, Ed., *Handbook of Green Chemistry*, Wiley-VCH, New York, 2010.

Ayres, Robert U. and Leslie W. Ayres, Eds., *A Handbook of Industrial Ecology*, Edward Elgar Publishing, Cheltenham, UK., 2002

Ayres, Robert U., and Benjamin Warr, *The Economic Growth Engine: How Energy and Work Drive Material Prosperity*, Edward Elgar, Northampton, MA, 2009.

Beer, Tom, and Alik Ismail-Zadeh, *Risk Science and Sustainability: Science for Reduction of Risk and Sustainable Development for Society*, Kluwer Academic, Boston, 2003.

Booth, Douglas E., *Hooked on Growth: Economic Addictions and the Environment*, Rowman &Littlefield Publishers, Lanham, MD, 2004.

Clark James and Duncan MacQuarrie, *Handbook of Green Chemistry and Technology*, Blackwell Science, Malden, MA, 2002.

Cote, Ray, James Tansey, and Ann Dale, Eds.,*Linking Industry and Ecology: A Question of Design*, University of British Columbia Press, Vancouver, 2006.

DeSimone, Joseph M., William Tumas. Eds., *Green Chemistry Using Liquid and Supercritical Carbon Dioxide*, New York : Oxford University Press, New York, 2003.

Doble, Mukesh, and Anil Kumar Kruthiventi, *Green Chemistry and Processes*, Elsevier, Amsterdam, 2007.

Doxsee, Kenneth M., and James E. Hutchison, *Green Organic Chemistry: Strategies, Tools, and Laboratory Experiments*, Thomson-Brooks/Cole, Monterey, CA, 2004.

El-Haggar, Salah M., *Sustainable Industrial Design and Waste Management: Cradle-to-Cradle forSustainable Development*, Elsevier Academic Press, Amsterdam, 2007.

Graedel, Thomas E. and Braden. R. Allenby, *Industrial Ecology and Sustainable Engineering*, Prentice Hall, Upper Saddle River, NJ, 2009.

Graedel, Thomas E. and Braden R. Allenby, *Industrial Ecology*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2003.

Graedel, Thomas E., and Jennifer A. Howard-Grenville, *Greening the Industrial Facility: Perspectives, Approaches, and Tools*, Springer, Berlin, 2005.

Grossman, Elizabeth.,*Chasing Molecules: Poisonous Products, Human Health, and the Promise of Green Chemistry*, Shearwater, DC, 2009.

Gupta, Surendra M.,and A.J.D. Lambert, Eds.*Environment Conscious Manufacturing*, Taylor &Francis/CRC Press, Boca Raton, FL, 2008.

Hawken, Paul, Amory Lovins, and L. Hunter Lovins, *Natural Capitalism: Creating the Next Industrial Revolution*, Back Bay Books, Boston, 2008.

Hendrickson, Chris T., Lester B. Lave, and H. Scott Matthews, *Environmental Life Cycle Assessment of Goods and Services: An Input-Output Approach*, Resources for the Future, Washington, DC, 2006.

Horvath, Istvan T., and Paul T. Anastas, “Innovations and Green Chemistry,” *Chemical Reviews*, **107**, 2169-2173 (2007).

Hunkeler, David, Kerstin Lichtenvort, and Gerald Rebitzer, Eds., *Environmental Life CycleCosting*, Taylor & Francis/CRC Press, Boca Raton, FL, 2008.

Islam, M. Rafiqul, Ed., *Nature Science and Sustainable Technology*, Nova Science Publishers, New York, 2007.Keeler, Marian, and Bill Burke, *Fundamentals of Integrated Design for Sustainable Building*, John Wiley & Sons, Hoboken, NJ, 2009.

Kronenberg, Jakub, *Ecological Economics and Industrial Ecology: A Case Study of the Integrated Product Policy of the European Union*, Routlege, New York, 2007.

Kutz, Myer, Ed., *Environmentally ConsciousTransportation*, Wiley, Hoboken, NJ, 2008.

Lankey, Rebecca L., and Paul T. Anastas, Eds., *Advancing Sustainability through Green Chemistry and Engineering*, American Chemical Society, Washington, DC, 2002.

Li, Chao-Jun, and Barry M. Trost, "Green Chemistry for Chemical Synthesis," *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 13197-13202 (2008).

Lifset, Reid, and Thomas E.Graedel, "Industrial Ecology: Goals and Definitions," in *AHandbook of Industrial Ecology*, Robert U. Ayres and L. Ayres, Eds., Edward Elgar, Cheltenham, UK, 2002.

Lutz, Wolfgang, and Warren Sanderson, Eds.,*The End of World Population Growth: Human Capital and Sustainable Development in the 21st Century*, Earthscan, Sterling, VA, 2004.

Matlack, Albert, *Introduction to Green Chemistry*, 2nd ed., Taylor & Francis/CRC Press, 2010.

McDonough, William, and Michael Braungart, *Cradle to Cradle: Remaking The Way We Make Things*, North Point Press, 2002.

Nelson, William M., *Green Solvents for Chemistry: Perspectives and Practice*, Oxford University Press, New York, 2003.

Nelson, William M., Ed., *Agricultural Applications in Green Chemistry*, Oxford University Press, New York, 2004.

Roesky, Herbert W. Dietmar Kennepohl, and Jean-Marie Lehn, Eds., *Experiments in Green and Sustainable Chemistry*, Wiley-VCH, New York, 2009.

Sheldon, Roger A., Isabel Arends, and Ulf Hanefield, *Green Chemistry and Catalysis*, Wiley-VCH, New York, 2007.

Simpson, R. David, Michael A. Toman, and Robert U. Ayres, Eds., *Scarcity and Growth Revisited: Natural Resources and the Environment in the New Millennium*, Resources for the Future, Washington, DC, 2005.

Tundo, Pietro, Alvise Perosa, and Fulvio Zecchini, *Methods and Reagents for Green Chemistry*, Wiley-Interscience, Hoboken, NJ, 2007.

Tundo, Pietro, and Vittorio Esposito, Eds.,*Green Chemical Reactions*, Springer, Dordrecht, Netherlands, 2008.

Vallero, Daniel A., *Sustainable Design: The Science of Sustainability and Green Engineering*, Wiley, Hoboken, NJ, 2008.

2.5.13.16: Literature Cited and Supplementary References is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Literature Cited and Supplementary References** has no license indicated.

2.5.13.17: Questions and Problems

1. Define industrial ecology.

2. Define an industrial ecosystem.

3. Name four major parts of an industrial ecosystem

4. Give the name of the processes to which materials and components are subjected in industrial ecosystems.

5. Give a definition of wastes in terms of natural resources.

6. What is the general pathway of materials through industrial systems as they currently operate?

7. What is meant by “level of recycling” and how is it related to embedded utility?

8. Name the three kinds of analyses and three categories considered in a life-cycle assessment.

9. What are the three kinds of products, classified in part on their amenability to recycling, that are normally considered in life-cycle assessments?

10. Give three important useful characteristics of consumable products related to their potential environmental effects.

11. Name three key attributes of industrial ecosystems that largely determine the wellbeing of the systems.

12. Given that an abundant source of energy can make almost anything possible in an industrial ecosystem, in what respects do vast reserves of coal, wind power, and solar energy fall short of being ideal energy sources?

13. Explain cogeneration of energy. What are its advantages?

14. Name three approaches to providing materials other than from virgin sources.

15. Consumable items and products cannot be recycled on a practical basis. Name three other categories of goods or products that can be recycled.

16. What is Kalundborg, Denmark, noted for?

17. Name several characteristics that facilitated development of the Kalundborg industrial ecosystem.

18. Name an environmental effect of industrial activities in the anthrosphere upon each of the other four environmental spheres.

19. Name an environmental effect of agricultural activities in the anthrosphere upon each of the other four environmental spheres.

20. For each of the other four environmental spheres, name a measure that may be taken in anthrospheric activities to reduce the environmental impact of these activities.

21. What does atom economy have to do with reducing the environmental impact of industrial ecosystems?

22. Describe two general approaches to making chemical synthesis greener.

23. What are structure activity relationships applied to toxic substances?

24. Why should the use of lipid-soluble organics be reduced in the practice of green chemistry?

25. What is a characteristic of VOCs that both makes them useful in industrial applications, but also increases their hazards?

26. Distinguish between reactive and corrosive substances. May a compound be both?

28. Distinguish among feedstocks, reagents, media, and catalysts.

29. Name two desirable characteristics of reagents insofar as their ability to generate products is concerned.

30. What are two oxidants used by organisms? What catalysts are used with these oxidants?

31. What is a nonchemical alternative to the use of oxidant and reductant reagents?

32. What are the two most common media used for chemical reactions?

33. Why is toluene preferred to benzene as a solvent medium for organic reactions?

34. What is a common toxicity problem with some lighter hydrocarbons, such as hexane?

35. Normally encountered as a gas, how can carbon dioxide be used as a medium for chemical reactions?

36. What are ionic liquids and what is their potential use in green chemistry?

37. Name two general categories of catalysts based upon physical form. Which is more desirable from the green chemistry standpoint?

38. What is a desirable characteristic of catalysts that chemists try to enhance?

2.5.13.17: Questions and Problems is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** has no license indicated.

SECTION OVERVIEW

2.5.14: Feeding the Anthrosphere- Utilizing Renewable and Biological Materials

"For approximately one century from the early 1900s to the early 2000s, petroleum-based feedstocks gave rise to a vast petrochemicals industry that resulted in the production of synthetic rubber, plastics, polymers with a variety of properties, pesticides and literally hundreds of other products, many of which replaced materials biosynthesized in nature, especially by plants. Now with diminishing petroleum supplies, a new generation of materials made from biomaterials such as the lignocellulose that composes plant structural matter is developing that will take the place of many of the petroleum-based chemicals. This massive shift to renewable feedstocks is leading to a new age of green chemistry."

2.5.14.1: Feeding the Anthrosphere

2.5.14.2: Utilization of Feedstocks

2.5.14.3: Biological Feedstocks

2.5.14.4: Biological Sources of Chemicals

2.5.14.5: Biorefineries and Biomass Utilization

2.5.14.6: Monosaccharide Feedstocks - Glucose, Fructose, and Xylose

2.5.14.7: Cellulose

2.5.14.8: Feedstocks from Cellulose Wastes

2.5.14.9: Lignin

2.5.14.10: Direct Biosynthesis of Polymers

2.5.14.11: Bioconversion Processes for Synthetic Chemicals

2.5.14.12: Bamboo - Ancient Material for the Future

2.5.14.13: Literature Cited

2.5.14.14: Questions and Problems

This page titled 2.5.14: Feeding the Anthrosphere- Utilizing Renewable and Biological Materials is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.14.1: Feeding the Anthrosphere

As the industrial revolution gained impetus since about 1800, and especially since about 1900 with the development of the chemical industry, the anthrosphere developed a voracious appetite for materials. This is especially true of the vast petrochemicals industry fed by materials from petroleum and producing huge quantities of polymers, plastics, synthetic detergents, agricultural chemicals and many other products. The era of petrochemicals must come to an end because sources of petroleum cannot sustain the enormous appetite of the anthrosphere for petrochemicals. The demand for the kinds of products now produced from petroleum will not go away, so alternate means of providing the materials desired by humans will have to be met. The only real alternative is biomaterials which in fact provided most of the stuff that humans used until very recently in the history of humankind. Although challenging, this shift in raw materials sources promises to be a very exciting one for chemistry. And it provides an opportunity for chemists and engineers to “get it right” by applying the principles of green chemistry, green engineering, and industrial ecology in ways that will ensure a sustainable future.

Feedstocks

Recall from Chapter 13, Section 13.15 that **feedstocks** are the main ingredients that go into the production of chemical products. Reagents act upon feedstocks and often the two are not readily distinguished. Feedstock selection largely dictates the reactions and conditions that will be employed in a chemical synthesis and is, therefore, of utmost importance in the practice of green chemistry. A feedstock should be as safe as possible. The source of a feedstock can largely determine its environmental impact, and the acquisition of the feedstock should not strain Earth’sresources. The process of isolating and concentrating a feedstock can add to the potential harm of otherwise safe materials. This is true of some metal ores in which corrosive and toxic reagents (in the case of gold, cyanide) are used to isolate the desired material.

As a general rule, it is best if feedstocks come from renewable sources rather than depletable resources A biomass feedstock, for example, can be obtained as a renewable resource grown by plants on land, whereas a petroleum-based feedstock is obtained from depletable crude oil resources. However, the environmental tradeoffs between these two sources may be more complex than first appears in that the petroleum feedstock may simply be pumped from a few wells in Saudi Arabia, whereas the biomass may require large areas of land, significant quantities of fertilizer, and large volumes of irrigation water for its production. Another important decision is whether or not the feedstock should be made entirely from virgin materials or at least in part from recycled material.

In the United States petroleum amounts to all but about two percent of the raw material used for the manufacture of organic chemicals and the many products made from them, such as textiles, plastics, and rubber. To a degree petroleum is an ideal feedstock for this purpose; during the last 100 years it has been readily available and relatively inexpensive except during times of temporary supply disruption. There are, of course, disadvantages to the use of petroleum as a feedstock, not the least of which is the fact that eventually available supplies will become exhausted. The transportation and refining of petroleum consume large amounts of energy, amounting to more than 15 percent of total energy use in United States. Chemically, a consideration with the use of petroleum as a raw material is that the hydrocarbon molecules that compose petroleum are in a highly reduced chemical state. In order to be utilized as feedstocks, petroleum hydrocarbons often must be oxidized. The oxidation process (see Section 13.15) entails a net consumption of energy and often requires the use of severe and hazardous reagents. Although commonly used oxidation processes are well contained and safe, there is always the consideration of possible combustion and explosion hazards in the partial oxidation of petroleum.

Much of the challenge and potential environmental harm in obtaining feedstocks is in separating the feedstock from other materials. This is certainly true with petroleum, which consists of many different hydrocarbons, only one of which may be needed as the raw material for a particular kind of product. Some metals occur at levels of less than 1% in their ores, requiring energy-intensive means of separating out the metals from huge quantities of rock. The smelting of copper and lead ores releases significant quantities of impurity arsenic with the flue dust, which must be collected from the smelting operation. Indeed, this byproduct arsenic provides more than enough of the arsenic needed in commerce. Biobased materials are also generally mixtures that require separation. Cellulose from wood, which can be converted to paper and a variety of chemicals, is mixed intimately with lignin, from which it is separated only with difficulty.

In evaluating the suitability of a feedstock, it is not sufficient to consider just the hazards attributable to the feedstock itself and its acquisition. That is because different feedstocks require different processing and synthetic operations downstream that may add to

their hazards. If feedstock A requires use of a particularly hazardous material to convert it to product, whereas feedstock B can be processed by relatively benign processes, feedstock B should be chosen. This kind of consideration points to the importance of considering the whole life cycle of materials rather than just one aspect of them.

This page titled 2.5.14.1: Feeding the Anthrosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **14.1: Feeding the Anthrosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.14.2: Utilization of Feedstocks

Before considering sources of feedstocks, it is useful to consider how those feedstocks can be used in the least polluting, most sustainable way possible. Feedstocks are modified by chemical processes to produce new chemical materials with commercial uses. The ideal feedstock is renewable and poses no hazards. And it can be converted to the desired product using few steps with 100% yield and 100% atom economy. This should be done with minimum quantities of reagent using only safe media in which the reaction occurs.

There are three major categories of reactions that are involved in chemical processing of feedstocks as shown in a general sense in Figure 14.1. In an **addition reaction**, all feedstock material becomes part of the product and there are no byproducts. These are the best kinds of reactions from the viewpoint of green chemistry because, when they work ideally, there are no wastes. A **substitution reaction** uses a reagent to replace a functional group on the feedstock molecule. As its name implies, an **eliminationreaction** removes a functional group from a feedstock molecule. Both of these latter kinds of reactions produce byproduct materials from the feedstock and from spent reagent. Their impacts can be reduced by reclaiming byproducts, if a use can be found for them, and by regenerating reagent, when that is possible. In some cases,elimination reactions can be carried out without use of a reagent, reducing the impact of this kind of reaction.

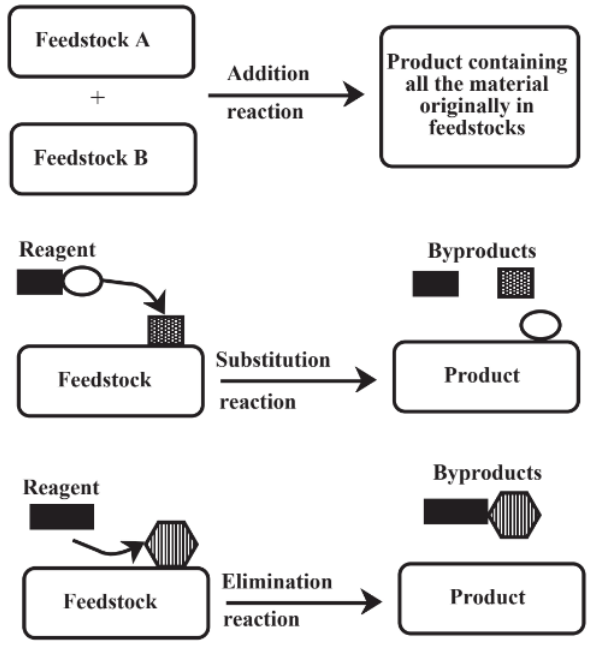


Figure 14.1. Illustration of three major categories of reaction processes by which feedstocks are acted upon by reagents to produce desired products

This page titled 2.5.14.2: Utilization of Feedstocks is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **14.2: Utilization of Feedstocks** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.14.3: Biological Feedstocks

Organisms have provided a huge share of the materials used by humans throughout their existence. Trees have served as sources of wood for construction and firewood. Animals provided hides and fur to protect primitive humans from Ice Age cold. The American bison was the source of food, shelter, clothing, and a wide variety of other useful items to plains-dwelling Native Americans. Protein silk is obtained from the cocoons of insects, and protein wool from sheep and related animals.

Biomass, the plant material generated from photosynthesis is the leading candidate to replace petroleum as a feedstock for the organic chemicals industry. There are several major categories of biomass that can be used for feedstock:

1. Carbohydrate, which has the general formula of approximately CH_2O. Carbohydrate is the biomass that is produced initially as glucose sugar from water and carbon dioxide during photosynthesis. It is contained in the structural parts of plants as cellulose, a biopolymer.

2. Lignin, a biological polymer with a complex structure, which occurs with carbohydrate cellulose in woody parts of plants, binding fibers of cellulose together. Relatively few uses have been found for lignin, and it poses impurity problems in extracting cellulose for feedstock use.

3. Lipid oils extracted from seeds, including soybeans, sunflowers, and corn.

4. Hydrocarbon terpenes produced by rubber trees, pine trees, and some other kinds of plants.

5. Proteins, produced in relatively small quantities, but potentially valuable as nutrients and other uses.

Biological materials used as sources of feedstocks are usually complex mixtures, which makes separation of desired materials difficult. However, in some biological starting materials nature has done much of the synthesis. Most biomass materials are partially oxidized as is the case with carbohydrates, which contain approximately one oxygen atom per carbon atom (compared to petroleum hydrocarbons which have no oxygen). This can avoid expensive, sometimes difficult oxidation steps, which may involve potentially hazardous reagents and conditions.

There are several main pathways by which feedstocks can be obtained from biomass. The most straightforward of these is a simple physical separation of biological materials, such as squeezing oil from oil-bearing biomass or tapping latex from rubber trees. Only slightly more drastic treatment consists of extraction of oils by organic solvents. Physical and chemical processes can be employed to remove useful biomass from the structural materials of plants, which consist of lignocellulose composed of cellulose and the related carbohydrate polymer hemicellulose bound together by lignin “glue.”

Carbohydrates are perhaps the most likely candidates as feedstocks for chemical processes. Carbohydrates come in several forms. Sucrose sugar, $C_{12}H_{22}O_{11}$, can be squeezed from sugar cane as sap and can be extracted from sugar beets and sugar cane with water. The exceptional photochemical productivity of sugar cane makes sucrose from this source an attractive option. Larger amounts of carbohydrates are available in starch, a polymer of glucose readily isolated from grains, such as corn, or from potatoes. An even greater source is found in cellulose, which occurs in woody parts of plants. It is relatively easy to break down starch molecules with the addition of water (hydrolysis) to give simple sugar glucose. Breaking down cellulose is more difficult, but can be accomplished by the action of cellulase enzymes.

Lipid oils are extracted from the seeds of some plants. Volatile solvents, most commonly the 6-carbon straight-chain alkane *n*-hexane, C_6H_{14}, are used to extract oils. In this process, the solvents are distilled off from the extract and recirculated through the process.

The hydrocarbon terpenes that occur in rubber trees can be tapped from the trees as a latex suspension in tree sap. Steam treatment and distillation can be employed to extract terpenes from sources such as pine or citrus tree biomass.

Grain seeds are rich sources of protein, almost always used for food, but potentially useful as chemical feedstocks for specialty applications. An exciting possibility just now coming to fruition in a practical sense is to transplant genes into plants so that they will make specialty proteins, such as medicinal agents.

This page titled 2.5.14.3: Biological Feedstocks is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

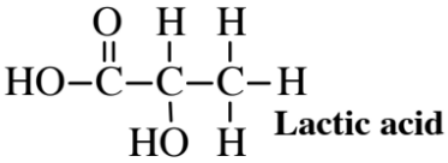
- **14.3: Biological Feedstocks** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.14.4: Biological Sources of Chemicals

In the provision of specialty and commodity chemicals and feedstocks, there are two main biological sources of materials. One of these consists of plants, which make huge quantities of cellulose and lesser quantities of other materials by photosynthesis. The other source is microorganisms, especially bacteria and yeasts.

Fermentation

Fermentation refers to the action of microorganisms on nutrients under controlled conditions to produce desired products. Fermentation for some products is anoxic (absence of O2); for other products oxic fermentation may be required. Fermentation processes have been used for thousands of years to produce alcoholic beverages, sauerkraut, vinegar, pickles, cheese, yogurt, and other foods. Ethanol, the alcohol in alcoholic beverages, is the most widely produced chemical made by fermentation. Lactic acid,



Lactic acid

has also been produced by fermentation processes for many years. More recently, fermentation has been applied to the production of a wide variety of organic acids, antibiotics, enzymes, and vitamins.

Starting in the 1940s, one of the major products of industrial fermentation has been penicillin, of which there are several forms. Figure 14.2 shows a simplified diagram of a facility for production of this life-saving antibiotic. Following penicillin, fermentation processes were developed for the production of several other significant antibiotics.

Selection of the appropriate microorganism is the most important consideration of a successful fermentation production process. The microorganisms have to have the proper nutrients, the choice of which can affect the kind and yield of the product. Sterile conditions must be maintained, and sterilization of equipment and media is accomplished by heating to 125–150°C for appropriate lengths of time. Air entering the fermenter must be filtered and sterilized. The temperature of fermentation is important, with fermentation rates increasing up to an optimum temperature, after which they decrease sharply with increased temperatures as the enzymes used by the microorganisms are destroyed (denatured). This kind of temperature relationship has increased interest in the use of thermophilic microorganisms that exist at boiling water temperatures in hot springs. If such organisms can be engineered to produce desired products, the rate of product generation may increase markedly. Both the levels of oxygen (which must be excluded from anoxic processes) and pH must be controlled precisely. Modern fermentation processes use a variety of sensors to continuously monitor conditions in the fermentation tank and computerized control to accurately control all the parameters.

Fermentation is undergoing tremendous development with the use of transgenic microorganisms to which genes have been transferred to make specific kinds of substances. The most common and valuable substances made by transgenic microorganisms consist of a variety of proteins. These include proteins and smaller molecule polypeptides that are used as pharmaceuticals. The best example of such a substance is human insulin, which is now produced in large quantities by transgenic microorganisms.

Until recently, fermentation has not been widely employed to make commodity chemicals used on a large scale. An exception is the large-scale production of ethanol from the fermentation of glucose sugar by yeasts. Now mandated as a gasoline additive in the U. S. by law, huge and growing quantities of ethanol are made by fermentation of glucose derived from corn and this use is an important market for corn. It is not clear that this is a truly green technology and some authorities believe that the energy consumed and the environmental damage from more intensive cultivation of corn outweigh the benefits of using this grain to produce ethanol fuel. Advances in transgenic microbiology have increased the possibilities for using fermentation to produce a variety of chemicals and chemical feedstocks, several examples of which are discussed in this chapter.

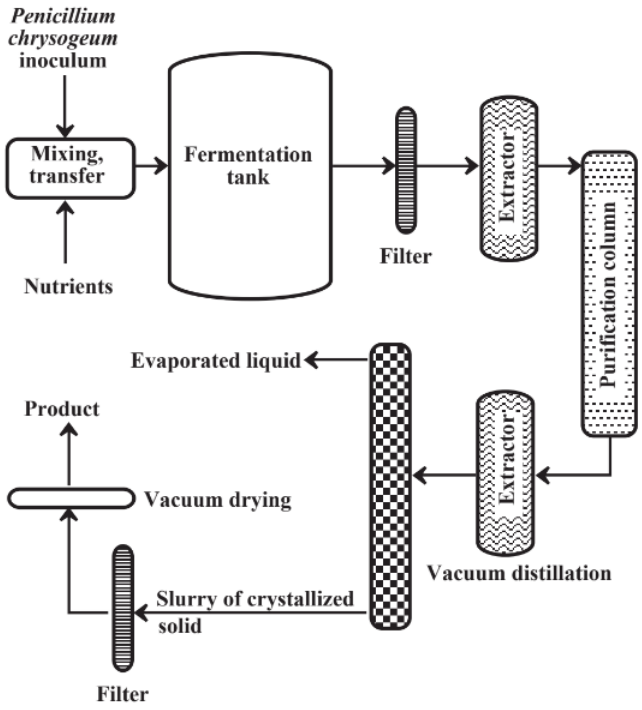


Figure 14.2. Simplified schematic diagram of the process for making penicillin by fermentation

Production of Materials by Plants

The uses of microorganisms operating in fermentation processes to generate commodity chemicals were discussed above. Plants are the other kind of organism that can be used for producing chemicals. Indeed, the nutrients used for fermentation processes come originally from plants. Fermentation is in a sense not a very efficient means of producing chemicals because of the consumption of nutrients to support the microorganisms and their reproduction and because of the generation of large quantities of byproducts. Plants, which generate their own biomass from atmospheric carbon dioxide and water are very efficient producers of materials. Wood and the cellulose extracted from it are prime examples of such materials.

In addition to their efficient production of biomass, plants offer distinct advantages in their production and harvesting. Genetics determine the materials that a plant makes, and once a crop is growing in a field, the products it is programmed for will be produced without fear of contamination by other organisms, which is always a consideration in fermentation. Plants can be grown by relatively untrained personnel using well known agricultural practices. Plant matter is generally easy to harvest in the form of grains, stalks, and leaves, which can be taken to a biorefinery (see below) to extract needed materials.

The production of feedstocks and other chemical commodities from plants has been limited by the genetic restrictions inherent to plants. Now, however, transgenic plants can be bred to produce a variety of materials directed by genes transplanted from other kinds of organisms. For example, as discussed in Section 14.10, plants have even been developed to synthesize plastics. Another limitation of the production of materials by plants has been the mixture of these materials with other matter generated by plants. The intimate mixture of useful wood cellulose with lignin, for which uses are still being sought, is a prime example of this problem. Again, transgenic technology can be expected to be helpful in developing plants that produce a relatively pure product (such as the almost pure cellulose in cotton.)

The potential of plants to produce useful products has been greatly increased by the development of hybrid plants with spectacular capacities to generate biomass by photosynthesis. Corn is one of the more productive field crops, and hybrid varieties produce large quantities of grain and plant biomass (leaves, stalks, husks and cobs commonly called corn stover). Sugar cane is noted for its ability to produce biomass, some in the form of sugar, much more in the cane stalk biomass. The sugar cane stalk residues left after extracting sugar from it (bagasse) has had relatively few uses, other than for fuel but potentially can produce large quantities of chemical feedstocks in biorefineries. One of the more exciting developments of productive hybrid plants is the hybrid poplar tree which, nourished by minimal amounts of fertilizer and watered by economical trickle irrigation systems, grows within a few years

to a harvestable size for the production of wood pulp and wood for plywood. The ability of these trees to generate cellulose that can be converted to glucose means that they may serve as the basis of an entire plant-based chemicals industry. The possibility exists that they can be genetically engineered to produce other chemicals as well.

This page titled 2.5.14.4: Biological Sources of Chemicals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **14.4: Biological Sources of Chemicals** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.14.5: Biorefineries and Biomass Utilization

Just as crude oil is a complex mixture of hydrocarbons and other organic compounds that must be run through a petroleum refinery to separate and chemically modify the materials in it to produce fuel and petrochemical feedstocks, chemicals from biological sources are usually complex mixtures that require refining and chemical processing to provide needed feedstocks. The separation and processing of biomaterials from plants or from bacteria cultured in digesters is accomplished in a **biorefinery** the output of which consists of organic chemicals and fuels. In a biorefinery the various products that can be obtained from biomass are separated and subjected to chemical treatment, generally with the objective of reducing the chemically bound oxygen contents of the organic liquids. Although the large number of compounds generated from biomass poses a challenge for separations, it also provides the opportunity to produce smaller quantities of high-value chemicals.

Figure 14.3 illustrates a biorefinery showing four ways in which feedstocks may be obtained from biomass. The simplest and least energy-consumptive of these is extraction in which an organic solvent or, in more sophisticated operations, supercritical carbon dioxide is used to dissolve materials from biomass. This approach is widely used to extract oils from some kinds of oil seeds. Terpene hydrocarbons can be extracted from pinewood and other terpene-producing plants. Some of the greatest potential is to extract hydrocarbons and other oils from algae that produce these materials. Single-cell algae are particularly efficient photosynthesizers and the potential is high to genetically engineer strains that can produce specific classes of extractable organics.

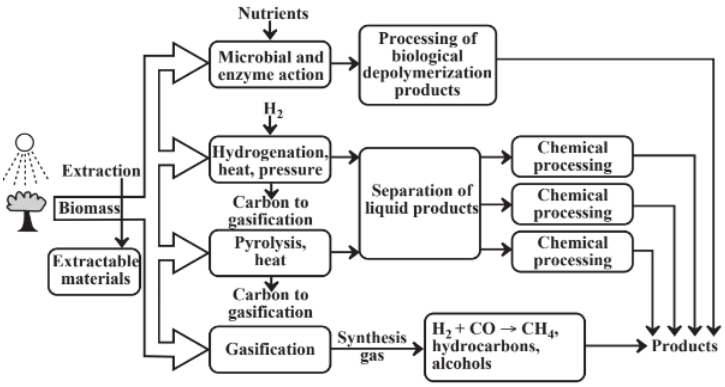


Figure 14.3. General outline of a biorefinery in which organic feedstocks and fuels are obtained by the action of microorganisms or enzymes that depolymerize large biomass molecules, extraction, hydrogenation, pyrolysis, or gasification of biomass.

Other than by extraction, the pathway to useful chemicals from biomass involves breaking down the complex biomass polymers. One way in which this is done is with microbial action or the action of isolated enzymes, for example, to produce glucose sugar from starch or cellulose. This step may require addition of nutrients including nitrogen, phosphorus, and potassium to enable microorganisms to grow. The glucose and other monomers isolated by enzymatic action can be subjected to additional processing, the most common example of which is fermentation of glucose to alcohol.

Hydrogenation of biomass involves reaction with elemental H_2 under high pressure and at elevated temperatures. This approach can be used with hydrogen generated relatively inexpensively by electrolysis of water employing renewable sources of electricity, especially from wind power. Direct hydrogenation of biomass produces a wide variety of organics including oxygenated compounds, some of which have direct uses and others of which may be chemically modified to give desired product.

Pyrolysis involves heating biomass externally or with a hot gas stream to evolve liquid and gas products. An external heat source including even solar energy focussed and concentrated on a reactor may be employed. As with hydrogenation, pyrolysis generates a variety of products including oxygenated compounds. It also produces large amounts of residual carbon, which can be used directly as fuel or gasified with steam and oxygen to produce synthesis gas.

Gasification, which is discussed in more detail in Chapter 15, Section 15.6, involves the reaction of hot carbon with steam,

$$H_2O + C \rightarrow H + CO_2$$

yielding a synthesis gas mixture of H_2 and CO. The carbon is usually heated by partial combustion with O_2 and the CO in the synthesis gas is reacted with steam to increase the ratio of H_2 to CO. Biomass can be gasified directly by reaction with a minimal

amount of O_2. Since biomass has the approximate empirical formula of $\{CH_2O\}$ the "water" required for gasification is largely in the biomass, itself. Such direct gasification of biomass also produces large quantities of organics that are processed downstream in the biorefinery.

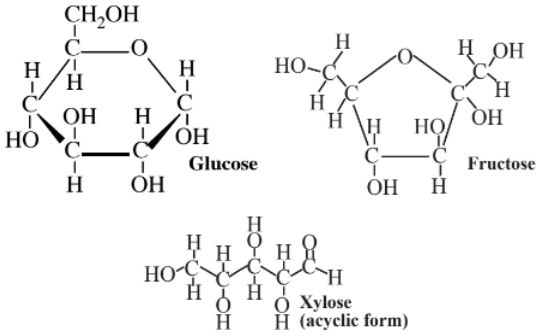
An important consideration in biorefineries is the use of catalysts. Insofar as possible biorefineries should use heterogeneous catalysts that do not get into the product and enzyme catalysts that operate at moderate temperatures.

This page titled 2.5.14.5: Biorefineries and Biomass Utilization is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **14.5: Biorefineries and Biomass Utilization** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

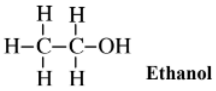
2.5.14.6: Monosaccharide Feedstocks - Glucose, Fructose, and Xylose

The monosaccharides such as glucose and fructose, as well as xylose, the monomer of hemicellulose, which makes up almost 1/3 of typical plant biomass,



are produced in abundance by plants. These compounds are excellent platforms for a number of different organic syntheses. As partially oxidized materials, they are particularly advantageous where a partially oxidized product is made, as is often the case in organic synthesis. Monosaccharides contain hydroxyl groups (-OH) around the molecule, which act as convenient sites for the attachment of various functionalities. Glucose is metabolized by essentially all organisms, so it serves as an excellent starting point for biosynthesis reactions using enzymes, and it and many of its products are biodegradable, adding to their environmental acceptability.

Glucose can be obtained by enzyme-catalyzed processes from other sugars, including sucrose and fructose. A large fraction of the glucose that is now used is obtained from the enzymatic hydrolysis of cornstarch. It is also possible to obtain glucose by the enzymatic hydrolysis of cellulose, although it has not proven economically practical to do so on an industrial scale because of the refractory nature of the cellulose polymer. Nevertheless, the enormous quantities of cellulose available in wood and other biomass sources make glucose from cellulose an attractive prospect. The greatest use of glucose and fructose (which is readily converted to glucose by enzymes) for synthesis is by fermentation with yeasts to produce ethanol,



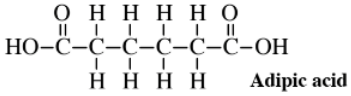
an alcohol widely used as a gasoline additive, solvent, and chemical feedstock. A byproduct of this fermentation process is carbon dioxide, the potential of which in green chemical applications as a supercritical fluid solvent are discussed in Section 13.14.

Glucose is widely used as a starting material for the biological synthesis of a number of different biochemical compounds. These include ascorbic acid, citric acid, and lactic acid. Several amino acids used as nutritional supplements, including lysine, phenylalanine, threonine, and tryptophan, are biochemically synthesized starting with glucose. The vitamins folic acid,ubiquinone, and enterochelin are also made biochemically from glucose.

In addition to the predominantly biochemical applications of glucose mentioned above, monosaccharides can be used to make feedstocks for chemical manufacture. The possibilities for so doing are now greatly increased by the availability of genetically engineered microorganisms that can be made to express genes for the biosynthesis of a number of products. Sophisticated genetic engineering is required to make chemical feedstocks because these are materials not ordinarily produced biologically.

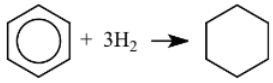
A study by the U.S. Department of Energy Pacific Northwest Regional Laboratory has identified “top twelve value added chemicals” that can be made enzymatically from monosaccharides, especially glucose and fructose.[1] Listed in Table 14.1, these chemicals could form the main feedstocks for future biorefineries that would generate an abundance of products currently made largely from petrochemicals. Several specific syntheses of commercially valuable chemicals are discussed below.

As an example of the potential of glucose for making important feedstocks, consider the synthesis from glucose of adipic acid,



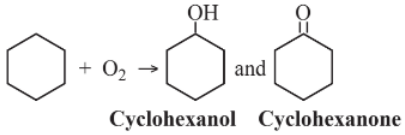
Adipic acid

a feedstock consumed in large quantities to make nylon. The conventional synthesis of this compound starts with benzene, a volatile, flammable hydrocarbon that is believed to cause leukemia in humans. The synthesis involves several steps using catalysts at high pressure and corrosive oxidant nitric acid, which releases air pollutant nitrous oxide, N_2O. The first step is the addition to benzene over a Ni/Al_2O_3 catalyst at a pressure 25 to 50 times atmospheric pressure of explosive hydrogen gas, H_2,



(14.6.1)

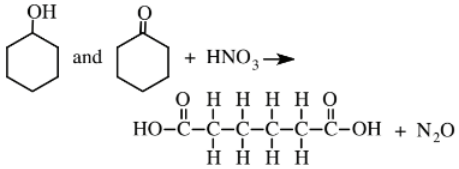
to produce cyclohexane, which is then subjected to oxidation in air at 9 atm pressure over a cobalt catalyst



Cyclohexanol **Cyclohexanone**

(14.6.2)

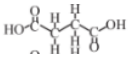
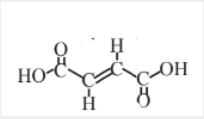
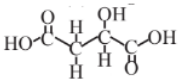
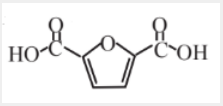
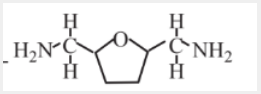
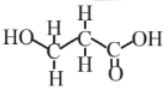
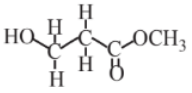
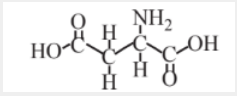
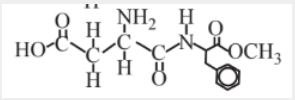
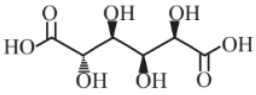
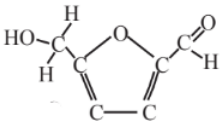
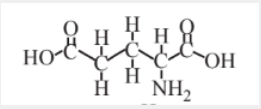
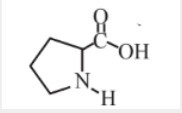
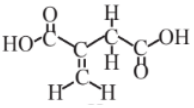
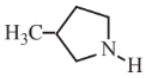
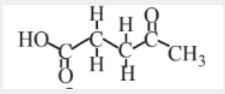
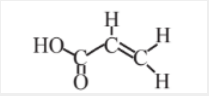
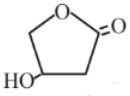
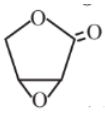
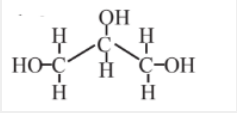
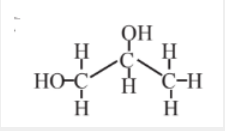
to produce a mixture of cyclohexanol, a cyclic alcohol, and cyclohexanone, a cyclic ketone. This mixture is then reacted with oxidizing, corrosive, 60% nitric acid over a Ni/Al_2O_3catalyst at 25–50 atm pressure to give the adipic acid feedstock:

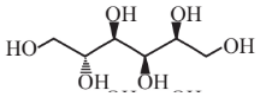
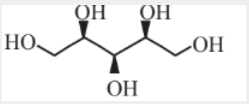
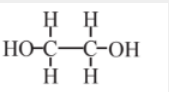


(14.6.3)

https://bio.libretexts.org/@go/page/108926

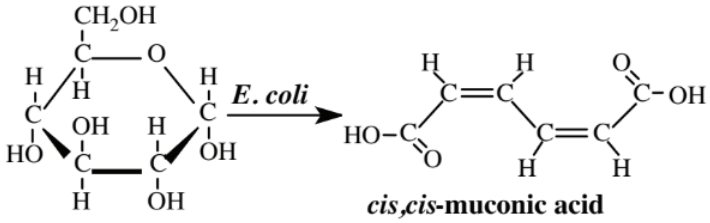
Table 14.1. Top Twelve Chemical Feedstock That Can be Made Enzymatically from Monosaccharide

Name and structural formula	Examples of Products
Four carbon 1,4-diacids	Tetrahydrofuran 
Succinic acid 	
Fumaric acid 	
Malic acid 	
2,5-Furandicarboxylic acid 	2,5-Bis(aminomethyl)-tetrahydrofuran 
3-Hydroxypropionic acid 	Methyl acrylate 
Aspartic acid 	Aspartame (artificial sweetener) 
Glucaric acid 	5-Hydroxymethyl-furfural 
Glutamic acid 	Proline 
Itaconic acid 	3-Methylpyrrolidine 
Levulinic acid 	Acrylic acid 
3-Hydroxybutyrolactone 	Epoxylactone 
Glycerol 	Propylene glycol 

Sorbitol 	Sorbitol itself has numerous uses in foods(as a low-calorie sweetener) and in cosmetics
Xylitol 	Ethylene glycol 

Throughout the synthesis process, elevated temperatures of approximately 250°C are employed. The N_2O released by the synthesis of adipic acid in the manufacture of nylon accounts for a significant fraction of worldwide N_2O releases. The potential dangers and environmental problems with this synthesis are obvious.

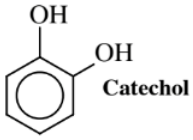
As an alternative to the chemical synthesis of adipic acid described above, a biological synthesis using genetically modified *Escherichiacoli* bacteria and a simple hydrogenation reaction has been devised. The bacteria convert glucose to *cis,cis*-muconic acid:



cis,cis-muconic acid (14.6.4)

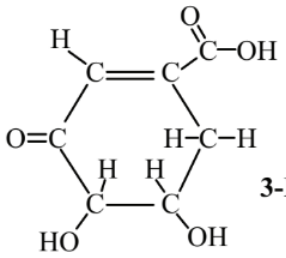
The muconic acid is then treated under relatively mild conditions with H_2 under 3 atm pressure over a platinum catalyst to give adipic acid.

Another organic chemical that potentially can be produced by the action of transgenic microorganisms on glucose is catechol, used as a feedstock to make flavors, pharmaceuticals, carbofuran pesticide, and other chemicals. About 20 million kilograms per year worldwide of this compound are now manufactured chemically starting with propylene and carcinogenic benzene, both derived from depleting petroleum sources. Toxic phenol is generated as an intermediate, and it is oxidized to catechol with 70% hydrogen peroxide, which at this concentration is a violently reactive, hazardous oxidant. These steps require some rather severe conditions and stringent precautions in handling hydrogen peroxide reagent.*E. coli* bacteria of a genetically modified strain designated AB2834/pKD136/pKD9/069A, produce catechol from glucose and, if yields can be gotten to acceptable levels, biosynthesis could become a major source of this important chemical.



Catechol

Another potentially important organic feedstock that has now been synthesized from glucose using transgenic *E. coli* is 3-dehydroshikimic acid:



3-Dehydroshikimic acid

This compound is an important intermediate in the production of aromatic amino acids, gallic acid, vanillin, and other chemicals. It also has antioxidant properties. Antioxidants are organic compounds that react with oxygen-containing, reactive free radical

species, such as hydroxyl radical, HO•. With their unpaired electrons (which make them free radicals), these species oxidize materials such as oils, fats, and lubricating oils and greases, causing deterioration in quality. By reacting with the free radicals, antioxidants stop their action. An abundant source of 3-dehydroshikimic acid could lead to its much wider application as an antioxidant.

The most abundant biomass feedstocks are carbohydrates. It follows that one of the most promising pathways to obtaining useful raw materials and fuels from biomass is their synthesis directly from carbohydrates. One of the more promising end products from chemical modification of carbohydrates is dimethylfuran, an oxygen-containing cyclic organic compound that has most of the desirable properties of hydrocarbons as a fuel and raw material. Compared to ethanol, dimethylfuran has a relatively low boiling temperature, has a high energy content per unit mass, does not absorb water, and exhibits combustion characteristics comparable to those of commonly used hydrocarbon fuels. Structurally, dimethylfuran with its 5-membered ring resembles the abundant monosaccharide fructose, which also has a 5-membered ring,

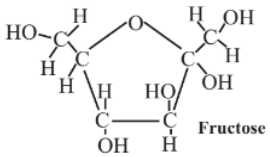


Figure 14.4.

This page titled 2.5.14.6: Monosaccharide Feedstocks - Glucose, Fructose, and Xylose is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **14.6: Monosaccharide Feedstocks - Glucose, Fructose, and Xylose** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.14.7: Cellulose

The most abundant natural material produced by organisms is **cellulose** synthesized biologically by the joining of glucose molecules with the loss of 1 H_2O molecule for each bond formed (see Figure 14.5). This makes the chemical formula of cellulose $(C_6H_{10}O_5)_n$, where n ranges from about 1500 to 6000 or more. Most cellulose is made by plants, with total amounts exceeding 500 *billion* metric tons per year world-wide. Cellulose makes up the sturdy cell walls of plants. Wood is about 40% cellulose, leaf fibers about 70%, and cotton, one of the purest sources of cellulose, about 95%. Cellulose occurs in different forms and is generally associated with hemicellulose (a material also composed of carbohydrate polymers) and lignin, a biopolymer of varied composition and bonding composed largely of aromatic unit.

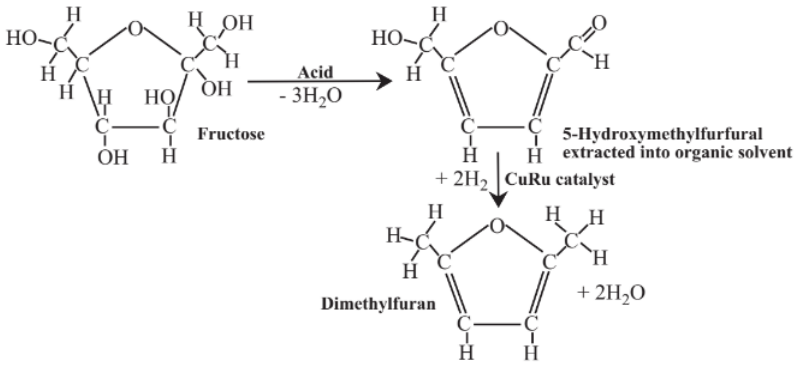


Figure 14.4. Synthesis of dimethylfuran, an organo-oxygen compound with many of the desirable properties of petroleum hydrocarbons, from fructose, a carbohydrate produced in abundance by plants.

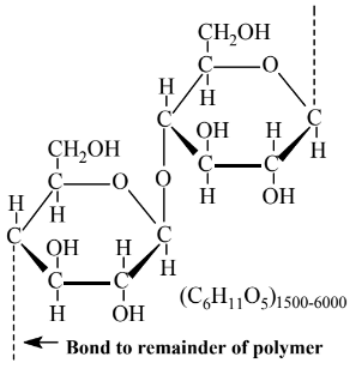


Figure 14.5. Segment of the cellulose molecule in which from 1500 to several thousand an hydroglucose units (glucose molecules less H_2O) are bonded together.

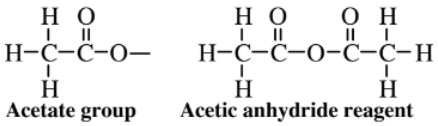
The first major step in cellulose utilization, such as extraction of cellulose fibers for making paper, consists of separating the cellulose from its matrix of lignocellulose (hemicellulose and lignin). This step has been the cause of many problems in utilizing cellulose because of the harsh chemical processing that has been employed. Lignin residues impart color to the cellulose, so wood pulp used in making paper has to be bleached with oxidants that alter the structure of the coloring agents. Bleaching used to be done almost entirely with elemental chlorine, Cl_2, and salts of hypochlorite ion, ClO^-. However, bleaching of biomass with these chlorine-based materials produces chlorinated organic impurities and pollutants. Therefore, ozone and hydrogen peroxide are preferred bleaching agents.

A finely divided form of cellulose called **microcrystallinecellulose** is produced by appropriate physical and chemical processing of cellulose. This material has many uses in foods in which it impart smoothness, stability, and a quality of thickness. Microcrystalline cellulose is also used in pharmaceutical preparations and cosmetics. Added to food, indigestible cellulose contributes bulk and retains moisture.

Chemically modified cellulose is used to make a wide variety of materials. Like the glucose that comprises it, cellulose has an abundance of -OH groups to which various other groups can be bonded to impart a variety of properties. One of the oldest synthetic fabrics, rayon, is made by treating cellulose with base and carbon disulfide, CS_2, then extruding the product through fine holes to

make thread. In a similar process, chemically treated cellulose is extruded through a long narrow slot to form a sheet of transparent film called cellophane.

As seen by the structural formula in Figure 14.5, each unit of the cellulose polymer has three–OH groups that are readily attached to other functional groups leading to chemically modified cellulose. One of the most common such products is cellulose acetate, an ester (see Section 6.4and Reaction 6.4.1) used primarily for apparel and home furnishings fabrics in which most of the -OH groups on cellulose are replaced by acetate groups by reaction with acetic anhydride (see below):



Although the cellulose feedstock for cellulose acetate synthesis is certainly a “green” material, acetic anhydride used to make the acetate is a corrosive, toxic chemical that produces poorly chealing wounds on exposed flesh. Furthermore, potentially hazardous solvents, such as dichloromethane, are used in some processes for making cellulose acetate.

Another cellulose ester that has been widely manufactured is cellulose nitrate in which the−OH groups on cellulose are replaced by $-ONO_2$ groups by treating cellulose with a mixture of nitric acid (HNO_3) and sulfuric acid (H_2SO_4). Cellulose nitrate makes transparent film and was used in the early days of moving pictures for movie film. However, one of the other major uses of this material is as an explosive, so cellulose nitrate can burn violently giving off highly toxic fumes of NO_2 gas. In years past this characteristic has lead to several tragic fires involving human fatalities. Its use is now largely restricted to lacquer coatings, explosives and propellants. Although the cellulose raw material is green, neither the process for making cellulose nitrate involving strong acids, nor the flammable product would qualify as green.

From the discussion above, it is apparent that cellulose is an important raw material for the preparation of a number of materials. The reagents and conditions used to convert cellulose to other products are in some cases rather severe. It may be anticipated that advances in the science of transgenic organisms will result in alternative biological technologies that will enable conversion of cellulose to a variety of products under relatively mild conditions.

This page titled 2.5.14.7: Cellulose is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **14.7: Cellulose** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.14.8: Feedstocks from Cellulose Wastes

Large quantities of cellulose-rich waste biomass are generated as byproducts of crop production in the form of straw remaining from grain harvest, bagasse residue from the extraction of sucrose from sugar cane, and other plant residues. This biomass represents a large amount of essentially free raw material that potentially can be converted to valuable chemical feedstocks. One way in which this can be done is by the use of enzyme systems to break the cellulose down into glucose sugar used directly as a feedstock or fermented to produce ethanol. Direct conversion of cellulose wastes to feedstocks is another route. Fortunately, nature has provided efficient microorganisms for this purpose in the form of rumen bacteria that live in the stomachs of cattle and related ruminant animals. It has been found that these bacteria function well in large fermenters from which oxygen is excluded if the plant residues are first treated with lime(chemical formulas $Ca(OH)_2$ and $CaCO_3$), producing short-chain organic acids that exist as their calcium salts in the presence of lime.

The organic acids produced by rumen bacteria in animals that can use cellulose as food are absorbed from the digestive systems of the animals and used as food. The acids produced in fermenters are in the form of calcium salts, the main ones of which consist of calcium acetate, calcium propionate, and calcium butyrate. These calcium salts of organic acids can be processed in several ways to produce feedstocks for a variety of organic syntheses. As shown in Figure 14.6, acidification of the salts yields the corresponding organic acids. Reaction of these acids with elemental hydrogen (hydrogenation) can be used to convert them to alcohols. Heat treatment(pyrolysis) of the calcium salts of the organic acids at 450°C produces ketones, such as those shown in Figure 14.6. These compounds are valuable feedstocks for a number of different chemical synthesis operations.

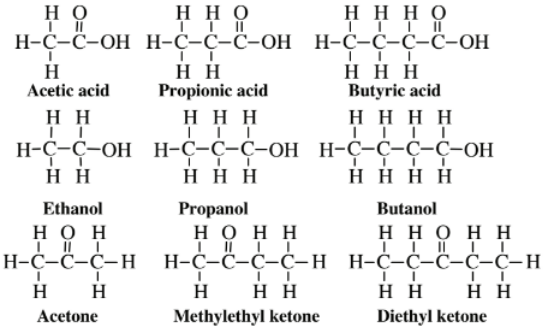


Figure 14.6. Production of organic chemicals from cellulose beginning with its breakdown by bacteria isolated from the rumen in animals capable of digesting cellulose. The first products are carboxylic acids, which can be converted to alcohols by hydrogenation and to ketones by pyrolysis

This page titled 2.5.14.8: Feedstocks from Cellulose Wastes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **14.8: Feedstocks from Cellulose Wastes** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.14.9: Lignin

Lignin, a chemically complex biopolymer that is associated with cellulose in plants and serves to bind cellulose in the plant structure, ranks second in abundance only to cellulose as a biomass material produced by plants. Lignin is normally regarded as a troublesome waste in the processing and utilization of cellulose. The characteristic that makes lignin so difficult to handle in chemical processing is its inconsistent, widely variable molecular structure as shown by the segment of lignin polymer in Figure 14.7. This structure shows that much of the carbon is present in aromatic rings that are bonded to oxygen-containing groups, and lignin is the only major plant biopolymer that is largely aromatic. Because of this characteristic, lignin is of considerable interest as a source of aromatic compounds including phenolic compounds, which have the−OH group bonded to aromatic rings or even aromatic hydrocarbons. The abundance of hydroxyl (−OH), methoxyl ($-OCH_3$), and carbonyl (C=O) groups in lignin also suggests potential chemical uses for the substance. A significant characteristic of lignin is its resistance to biological attack. This property, combined with lignin's highly heterogeneous nature makes it a difficult substrate to use for the enzyme-catalyzed reactions favored in the practice of green chemistry to give single pure products useful as chemical feedstocks.

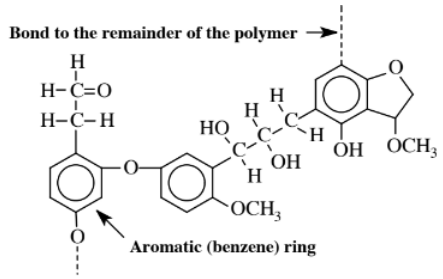


Figure 14.7. Segment of a lignin polymer molecule showing aromatic character and the disorganized, variable chemical structure that makes lignin a difficult material to use as a feedstock

Since lignin is a significant fraction of all plant biomass, significant fractions of this biopolymer must be dealt with in biorefineries. Lignin generated as a byproduct in the extraction of cellulose from wood is now largely burned for fuel, the lowest level of use for this material. By retaining much of the lignin molecule intact, use may be made of larger molecular mass segments of the molecule, such has been done for some uses for binders to hold materials together incoherent masses, fillers, resin extenders, and dispersants. There is also some potential to use lignin as a degradation-resistant structural material, such as in circuit boards. Potentially the most profitable use for lignin is to make small aromatic molecules useful for chemical synthesis. For this to be practical, special techniques will need to be developed to partially break down the lignin molecule without destroying the aromatic molecule segments in it.

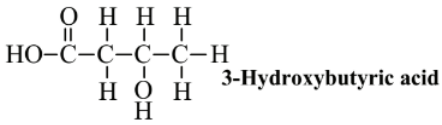
This page titled 2.5.14.9: Lignin is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **14.9: Lignin** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

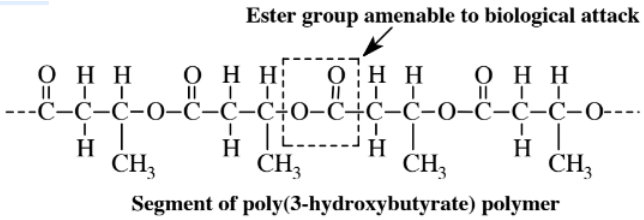
2.5.14.10: Direct Biosynthesis of Polymers

Cellulose in wood and cotton is only one example of the numerous significant polymers that are made biologically by organisms. Other important examples are wool and silk, which are protein polymers. A big advantage of these kinds of polymers from an environmental viewpoint is that polymers made biologically are also the ones that are most likely to be biodegradable. Attempts have been made to synthesize synthetic polymers that are biodegradable, These efforts have centered on those prepared from biodegradable monomers, such as lactic acid.

From the standpoint of green chemistry, it is ideal to have polymers that are made by organisms in a form that is essentially ready to use. Recently, interest has focussed on poly (hydroxyalkanoate) compounds, of which the most common are polymers of 3-hydroxybutyricacid:



This compound and related ones have both a carboxylic acid (-CO_2H) and an alcohol (-OH) group. As discussed in Section 6.4 and shown in Reaction 6.4.1, a carboxylic acid can bond with an alcohol with the elimination of a molecule of H_2O forming an *ester linkage*. Since the hydroxyalkanoates have both functional groups, the molecules can bond with each other to form polymer chains:



Segment of poly(3-hydroxybutyrate) polymer

Ester groups are among the most common in a variety of biological compounds, such as fats and oils, and organisms possess enzyme systems that readily attack ester linkages. Therefore, the poly(hydroxyalkanoate) compounds are amenable to biological attack. Aside from their biodegradability, polymers of 3-hydroxybutyric acid and related organic acids that have -OH groups on their hydrocarbon chains (alkanoates) can be engineered to have a variety of properties ranging from rubberlike to hard solid materials.

It was first shown in 1923 that some kinds of bacteria make and store poly(hydroxy-alkanoate) ester polymers as a reserve of food and energy. In the early 1980s it was shown that these materials have thermoplastic properties, meaning that they melt when heated and resolidify when cooled. This kind of plastic can be very useful, and the thermoplastic property is rare in biological materials. One commercial operation was set up for the biological synthesis of a polymer in which 3-hydroxybutyrate groups alternate with 3-hydroxyvalerate groups, where valeric acid has a 5-carbon atom chain. This process uses a bacterium called *Ralstoniaeutropia* fed glucose and the sodium salt of propionic acid (structural formula in Figure 14.6) to make the polymer in fermentation vats. Although the process works, costs are high because of problems common to most microbial fermentation synthesis processes: The bacteria have to be provided with a source of food, yields are relatively low, and it is difficult to isolate the product from the fermentation mixture.

Developments in genetic engineering have raised the possibility of producing poly(hydroxyalkanoate) polymers in plants. The plant *Arabidopsisthaliana* has accepted genes from bacterial *Alcaligeneseutrophus* that have resulted in plant leaves containing as much as 14% poly(hydroxybutyric acid) on a dry weight basis. Transgenic *Arabidopsisthaliana* and *Brassica napus* (canola) have shown production of the copolymer of 3-hydroxybutyrate and 3-hydroxyvalerate. If yields can be raised to acceptable levels, plant-synthesized poly(hydroxyalkanoate) materials would represent a tremendous advance in biosynthesis of polymers because of the ability of photosynthesis to provide the raw materials used to make the polymers.

This page titled 2.5.14.10: Direct Biosynthesis of Polymers is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

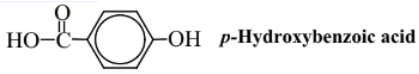
- **14.10: Direct Biosynthesis of Polymers** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.14.11: Bioconversion Processes for Synthetic Chemicals

Most of the biochemical operations described so far in this chapter pertain to natural products which, by their nature, would be expected to be amenable to the action of enzymes. The mild conditions under which enzymes operate, the readily available, safe reagents that they employ, such as molecular O_2 for oxidations, and the high specificity of enzyme catalysts make biocatalyzed reactions attractive for carrying out chemical processes on synthetic chemicals, such as those from petroleum sources. This section discusses two examples of enzyme-catalyzed processes applied to chemical processes on synthetic chemicals that would otherwise have to be performed with chemical reagents under much more severe conditions.

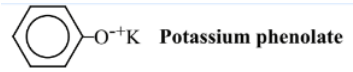
p-Hydroxybenzoic Acid from Toluene

The potential for use of biosynthesis applied to synthetic chemicals can be illustrated by the synthesis of *p*-hydroxybenzoic acid



***p*-Hydroxybenzoic acid**

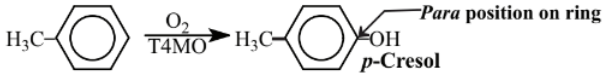
an important intermediate used in the synthesis of pharmaceuticals, pesticides, dyes, preservatives, and liquid crystal polymers currently made by reacting potassium phenolate,



Potassium phenolate

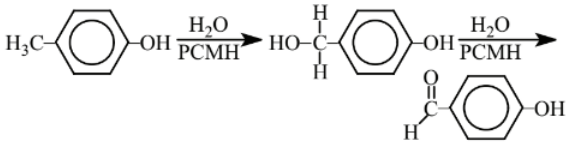
with carbon dioxide under high pressure at 220°C, which converts slightly less than half of the potassium phenolate to the desired product and produces substantial impurities. The process dates back to the early 1860s almost 150 years ago, long before there were any considerations of pollutants and wastes. It requires severe conditions and produces metal and phenol wastes. Reactive alumina powder (Al_2O_3) used to catalyze the process has been implicated in a 1995 explosion at a facility to produce *p*-hydroxybenzoic acid that killed 4 workers.

A biosynthetic alternative to the synthesis described above has been attempted with *Pseudomonasputida* bacteria genetically engineered to carry out several steps in the synthesis of *p*-hydroxybenzoic acid starting with toluene. A key to the process is the attachment at the *para* position on toluene of a hydroxyl group by the action of toluene-4-monooxygenase (T4MO) enzyme system transferred to *Pseudomonasputida* from *Pseudomonas mendocina*:



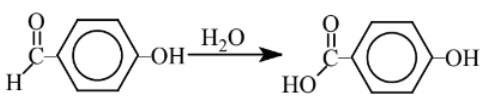
(14.10.1)

The next step is carried out by *p*-cresol methylhydroxylase (PCMH) enzyme from a strain of *Pseudomonasputida* that yields *p*-hydroxybenzyl alcohol followed by conversion to *p*-hydroxybenzaldehyde:



(14.10.2)

The last step is carried out by an aromatic aldehyde dehydrogenase enzyme designated PHBZ also obtained from a strain of *Pseudomonas putida* and consists of the conversion of the aldehyde to the *p*-hydroxybenzoic acid product:



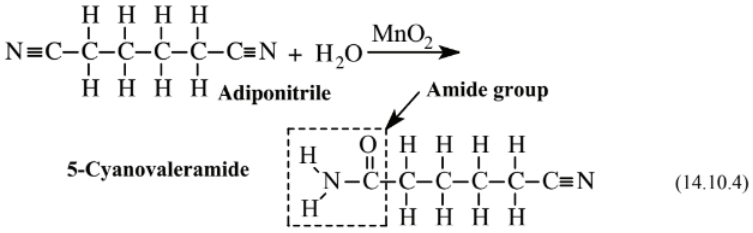
(14.10.3)

Through elegant genetic manipulation, the chemical processes described above were achieved leading to the desired product. In addition to providing the enzymes to carry out the desired steps, it was also crucial to block steps that would consume intermediates and give undesired byproducts that would consume raw material and require separation from the product. Although it is a long way

from showing that the complex biochemical synthesis process actually gives the desired product to the final goal of having a practical process that can be used on a large scale, the results described above certainly show the promise of transgenic organisms in carrying out chemical syntheses.

Production of **5-Cyanovaleramide**

The second biocatalyzed process to be considered is the conversion of adiponitrile to 5-cyanovaleramide. This conversion was required for the synthesis of a new chemical used for crop protection. This process can be carried out chemically with a stochichiometric mixture of adiponitrile with water and a manganese dioxide catalyst under pressure at 130° C as shown by the following reaction:



(14.10.4)

If the reaction is run to 25% completion, an 80% selectivity for the 5-cyanovaleramide is achieved, with the other fraction of the adiponitrile that reacts going to adipamide, in which the second C≡N functional group is converted to an amide group. Carrying the reaction beyond 25%completion resulted in unacceptable levels of conversion to byproduct adipamide.

The isolation of the 5-cyanovaleramide product from the chemical synthesis described above entails dissolving the hot reaction mixture in toluene solvent, which is then cooled to precipitate the product. The unreacted adiponitrile remains in toluene solution from which it is recovered to recycle back through the reaction. For each kilogram of 5-cyanovaleramide product isolated, approximately 1.25 kg of MnO_2 required disposal; this is definitely not a green chemical process!

As an alternative to the chemical synthesis described above, a biochemical synthesis was developed using organisms that had nitrile hydratase enzymes to convert the C≡N functional group to the amide group.[2] The microorganism chosen for this conversion was designated *Pseudomonaschloroaphis* B23. The cells of this organism were immobilized in beads of calcium alginate, the salt of alginic acid isolated from the cell walls of kelp. It was necessary to run the process at 5° C above which temperature the enzyme lost its activity. With this restriction, multiple runs were performed to convert adiponitrile to 5-cyanovaleramide. During these runs, 97% of the adiponitrile was reacted with only 4% of the reaction going to produce byproduct adipamide. The water-based reaction mixture was simply separated mechanically from the calcium alginate beads containing the microorganisms, which were then recycled for the next batch of reactant. The water was distilled off of the product to leave an oil from which the 5-cyanovaleramide product was dissolved in methanol, leaving adipamide and other byproducts behind. In contrast to the enormous amount of waste catalyst produced in the chemical synthesis of 5-cyanovaleramide, only 0.006 kg of catalyst waste residue was produced per kg of product. The waste microbial catalyst was 93% water, so its disposal was not a problem.

2.5.14.11: Bioconversion Processes for Synthetic Chemicals is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **14.11: Bioconversion Processes for Synthetic Chemicals** has no license indicated.

2.5.14.12: Bamboo - Ancient Material for the Future

Bamboo is a term given to large, fast-growing woody grasses including 1250 species within 75 genera ranging from small grassy plants that can substitute for lawn grass to giant bamboo or tree bamboo with stems 30 cm thick and comparable to trees in size. Bamboo has been used since ancient times and has approximately 1,500 applications, especially in Asia. These include structural and reinforcing fiber materials, paper, textiles, food, and fuel. Bamboo is a very strong material with a tensile strength comparable to that of steel. Bamboo textiles are now being used as substitutes for cotton. Although the separation from bamboo of fibers suitable for fabrics is a challenge, the fact that bamboo produces these fibers at a rate several times that of cotton makes bamboo fabrics very attractive as a cotton substitute.

Bamboo propagates primarily by rhizomes, underground stems that grow horizontally and that produce shoots and roots as they spread (good for reproduction but not desirable for neighboring areas where bamboo growth is not wanted). After harvesting, bamboo re-grows from its underground rhizome structure. Bamboo plants typically increase in biomass by 10-30% per year compared to 2-5% annually for trees. The annual yield of bamboo wood is about twice that of loblolly pine, a tree noted for its high productivity. The growth rate of some bamboo varieties under favorable conditions can be spectacular, exceeding 20 meters per year in some cases. Bamboo stalks sequester significantly more carbon dioxide and release correspondingly greater amounts of oxygen compared to trees of the same size.

Bamboo almost certainly has a bright future for sustainability as an attractive renewable source of biomass for materials and energy. There are several reasons for this related to the unique and diverse properties of bamboo, its high productivity, and the relatively short cycles over which bamboo can be produced. The production cycle of bamboo is generally much shorter than that of trees and the bamboo is generally harvested after 5-7 years of growth. Arguably, the greatest contribution that bamboo can make is for erosion control because of the dense underground soil-anchoring rhizome systems possessed by some of the more prominent bamboo varieties. Haiti, especially, could benefit from widespread growth of bamboo on eroded slopes denuded of once abundant forests to provide for firewood for cooking.

This page titled 2.5.14.12: Bamboo - Ancient Material for the Future is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **14.12: Bamboo - Ancient Material for the Future** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.14.13: Literature Cited

1. Werpy, T. and G. Petersen, *Top Value Added Chemicals from Biomass*, U.S. Department of Energy Pacific Northwest Regional Laboratory, Richland, WA, 2004, Vol. 1, available from www.osti.gov/bridge(Vol. 2 on lignin 2007).

2. DiCosimo, Robert, “Biocatalytic Production of 5 Cyanovaleramide from Adiponitrile,”*Green Chemical Syntheses and* Processes, Oxford University Press, New York, 2000, pp.114–125.

2.5.14.13: Literature Cited is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Literature Cited** has no license indicated.

2.5.14.14: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. Discuss advantages that biological feedstocks have over petroleum. Are there disadvantages?

2. What is a fundamental chemical difference between petroleum and biological feedstocks?

3. Potassium permanganate, sodium dichromate, and hydrogen peroxide may all three be used as reagents to react with organic molecules. What kind of reaction to they cause to occur? Which of these would be regarded as the greenest reagent and why is that so?

4. Name three kinds of reactions used in processing feedstocks. Which is best from the viewpoint of green chemistry?

5. Name several fractions of biomass that can be used for feedstocks. Which of these is the least useful?

6. How are oils extracted from plant sources?

7. Use chemical formulas to make the argument that carbohydrates are a more oxidized chemical feedstock than hydrocarbons.

8. Before the petrochemicals industry developed, the use of coal in making steel gave rise to a major organic chemicals industry? Explain why this was so and list some of the products that were made

9. Name some categories of chemicals routinely produced by fermentation.

10. Which pharmaceutical material has been produced by fermentation for many years?

11. What is the first, most important consideration in developing a fermentation process for production of a chemical?

12. What is the significance of temperature in fermentation processes? What happens if temperature is too high?

13. Which chemical is made in largest quantities by fermentation?

14. In which fundamental respect are plants more efficient producers of material than fermentation?

15. Which relatively recent advance in biotechnology has greatly increased the scope of materials potentially produced by plants?

16. Why are hybrid poplar trees particularly important in the production of raw materials?

17. Describe the structural characteristics of glucose and other carbohydrates that make them good platforms for chemical synthesis.

18. Give a disadvantage and an advantage of the use of cellulose as a source of glucose.

19. List some of the hazards associated with the chemical synthesis of adipic acid used to make nylon.

20. Give a major concern with the use of benzene as a feedstock.

21. What is a chemical characteristic of 3-dehydroshikimic acid that could lead to much greater uses for it?

22. Although the chemical formula of glucose is $C_6H_{12}O_6$, that of the cellulose polymer made from glucose is $(C_6H_{10}O_5)_n$ where n is a large number. Since cellulose is made from glucose, why is the cellulose formula not $(C_6H_{12}O_6)_n$?

23. Why is wood pulp consisting mostly of cellulose, treated with oxidants? Which oxidants are preferred, and which has lost favor?

24. Give some examples of useful chemically modified cellulose. Which of these has proven to be rather dangerous?

25. In ruminant animals that have bacteria in their stomachs that digest cellulose, the rumen bacteria and the organic acids they generate are passed on through the digestive tract where the bacterial biomass is dissolved, with the products and the organic acids previously generated absorbed by the animal as food. Suggest why basic limestone is used in the large batch processes that use rumen bacteria in digesters to produce organic acids from cellulose.

26. Why is it difficult to deal with lignin as a source of chemical feedstocks? What is the current main use of waste lignin?

27. Using internet resources, try to find an example of an operational biorefinery that is either in commercial operation or at a pilot plant level. What materials does it process and how does it process them? What are its main products?

28. Give the main advantage of biopolymers from an environmental viewpoint.

29. Which structural feature of hydroxyalkanoates enables them to make polymeric molecules?

30. What was the original source of poly(hydroxyalkanoate) polymers? How is it now proposed to produce them?

31. Although enzymes have not developed specifically to act upon synthetic compounds, they have some specific advantages that make them attractive for carrying out chemical processes on synthetic compounds. What are some of these advantages?

32. Name two chemicals for which it has been shown that enzymatic processes can actually convert synthetic raw materials to chemical products normally made by nonbiological chemical reactions.

2.5.14.14: Questions and Problems is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** has no license indicated.

SECTION OVERVIEW

2.5.15: Sustainable Energy- The Essential Basis of Green Systems

“If enough energy is available by means that are sustainable and that do not end up ruining theenvironment, virtually all sustainability needs may be met.”

2.5.15.1: Sustainability and Energy

2.5.15.2: What is Energy?

2.5.15.3: Radiant Energy from the Sun

2.5.15.4: Sources of Energy Used in the Anthrosphere

2.5.15.5: Conversions Between Forms of Energy

2.5.15.6: Green Technology for Energy Conversion

2.5.15.7: Energy Conservation

2.5.15.8: Depletable Fossil Fuels

2.5.15.9: Carbon Sequestration for Fossil Fuel Utilization

2.5.15.10: Nuclear Energy

2.5.15.11: Renewable Energy Sources - Solar Energy

2.5.15.12: Energy from Wind and Water

2.5.15.13: Biomass Energy

2.5.15.14: Geothermal Energy

2.5.15.15: Hydrogen for Energy Storage and Utilization

2.5.15.16: Literature Cited and Supplementary References

2.5.15.17: Questions and Problems

This page titled 2.5.15: Sustainable Energy- The Essential Basis of Green Systems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.15.1: Sustainability and Energy

Excluding the social and institutional problems of sustainability (as examples those of overpopulation or of vested economic interests that are environmentally counterproductive), sustainability issues are overwhelmingly energy issues. Most environmental and sustainability problems can be solved if abundant sources of energy are available and if they can be used without doing unacceptable harm to the environment. Consider the following:

•Material resources: With abundant energy, materials can be obtained in adequate quantities from sources normally considered to be marginal, such as metals from low-grade ores and organics from unconventional sources (even carbon dioxide can be converted to organics if energy is not an issue).

•Water: Earth has large quantities of water that is not suitable for domestic and other uses. If enough energy is available, saline ocean and ground water can be desalinated by distillation or reverse osmosis and wastewater can be purified to drinking water standards.

• Wastes: With enough energy the volume of municipal wastes can be greatly reduced and converted to a harmless mineral residue. Hazardous wastes can be made nonhazardous, their volume greatly reduced, and placed where they will do no harm.

•Food: Abundant energy enables production of adequate amounts of food. Water desalination, wastewater purification, and pumping water for long distances can provide needed irrigation water. Marginal land can be leveled and terraced and rocks crushed to a size that permits cultivation. High-value specialty foods can be grown inheated greenhouses even during winter.

•Transportation: Sustainable transportation is all about energy and can be achieved with energy-efficient vehicles and electrified rail systems.

•Fuels: Concentrated forms of carbon including biomass and even carbon dioxide can be converted to hydrocarbon fuels for applications for which there are no viable alternatives (such as aircraft) without adding any net amounts of greenhouse-gas carbon dioxide to the atmosphere.

• Dwellings and workplaces: The places where humans live and work can be maintained at comfortable temperatures regardless of harsh conditions outside.

Energy supply as such is not the problem; for example, world coal resources can meet energy needs for several centuries. The problem is that current patterns of energy use are not sustainable. Earth's peak petroleum production has probably already been reached. Using the remaining petroleum and switching to coal will almost certainly lead to global warming that might destroy Earth as we know it. The great challenge, therefore, is to develop sources of energy that are renewable and sustainable and that do not significantly damage Earth and its environment. A top priority must remain conservation of energy and greatly increased efficiency of energy utilization. Energy alternatives include wind, solar, biomass, geothermal, and nuclear energy sources along with some minor sources such as tidal energy. The use of fossil fuels will not cease and may continue for a long time with sequestration of carbon dioxide from combustion. This chapter discusses the energy alternatives listed above with emphasis upon energy sustainability.

This page titled 2.5.15.1: Sustainability and Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.1: Sustainability and Energy** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.2: What is Energy?

Energy is the ability to move matter around, that is, to do work. The movement of atoms and molecules is also a form of energy called **heat**. The energy contained in a moving mass of matter is **kinetic energy**. For example, energy collected from sunlight during the day can be accumulated in rapidly rotating spinning flywheels, then used at night when solar energy is not available. Water pumped into an elevated reservoir is an example of **potentialenergy** that can be run through a hydroelectric turbine to generate electricity as needed.

Chemical energy is a form of potential energy stored in the bonds of molecules. This energy can be released during chemical reactions, usually as heat but sometimes as electrical or light energy, as bonds are broken and new bonds are formed.

A crawler tractor equipped with a bulldozer for earth-moving illustrates the definition of energy and several forms of energy (Figure 15.1). **Chemical energy** in the form of petroleum hydrocarbons is used to fuel the tractor's diesel engine. In the engine the hydrocarbons combine with oxygen from air,

$$2\,C_{16}H_{34} + 49\,O_2 \rightarrow 32\,CO_2 + 34\,H_2O + \text{heat energy}$$

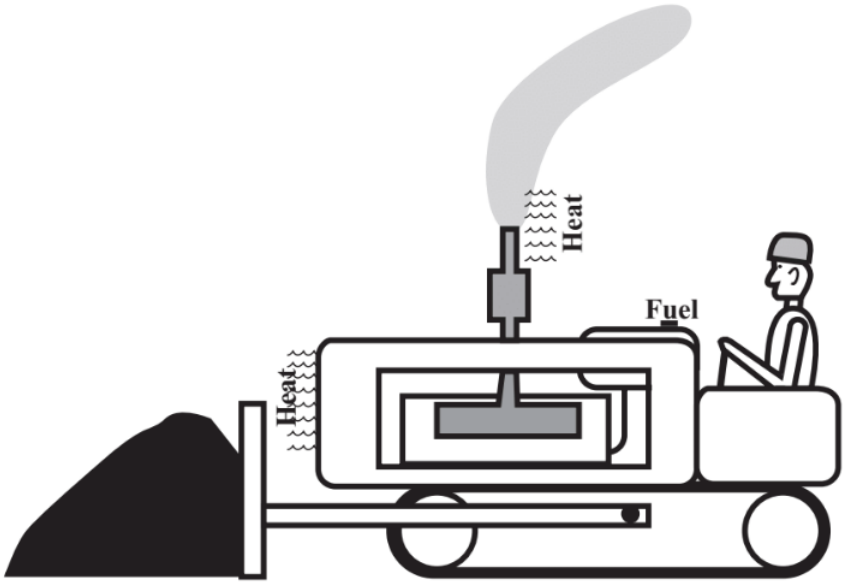


Figure 15.1. Conversion of chemical energy in diesel fuel to mechanical energy in an earth-moving operation through the intermediate production of heat energy. More than half of the chemical energy released in the combustion of the fuel in the tractor's engine is dissipated as heat dispersed through the engine's exhaust and cooling system

to produce **heat energy**. As the hot gases in the engine's cylinders push the pistons down, some of this heat energy is converted to **mechanical energy**, which is transferred by the engine crankshaft, gears, axle, and tracks to propel the tractor forward. A blade or other implement attached to the tractor moves soil.

The energy released in a chemical reaction results from the difference in energies between the bonds in the reactants and the energy of the bonds in the products. An example calculation is shown for the combustion of a mole of methane in Chapter 5, Section 9.5.

Energy Units and Thermodynamics

The standard unit of energy is the **joule**, abbreviated **J**. A total of 4.184 J of heat energy will raise the temperature of 1 g of liquid water by 1°C. This amount of heat is equal to 1 **calorie** of energy (1 cal = 4.184 J), the unit of energy formerly used in scientific work. A joule is a small unit, and the kilojoule, kJ, equal to 1000 J is widely used in describing chemical processes. The"calorie" commonly used to express the energy value of food (and its potential to produce fat) is actually a kilocalorie, kcal, equal to 1000 cal.

Power refers to energy generated, transmitted, or used per unit time. The unit of power is the **watt** equal to an energy flux of 1 joule per second (J s-1). A compact fluorescent light bulb adequate to illuminate a desk area might have a rating of 21 watts. A large powerplant may put out electricity at a power level of 1000 **megawatts** (mw, where one mw is equal to 1 million watts). Power on a national or global scale is often expressed in **gigawatts**, each one of which is equal to a billion watts or even **terawatts**, where a terawatt is equal to a trillion watts.

The science that deals with energy in its various forms and with work is **thermodynamics**. There are some important laws of thermodynamics. The **first law of thermodynamics** states that energy is neither created nor destroyed. This law is also known as the **law of conservation of energy**. As an example of the application of this law, consider Figure 15.1. The energy associated with moving earth enters the system as chemical energy in the form of diesel fuel, and the oxygen from the air required for its combustion. This is a valuable form of concentrated chemical energy that can be used to propel a tractor or locomotive, in a turbine attached to a generator for the generation of electrical energy, or as a fuel to generate heat in an oil-fired furnace. The fuel is burned in the tractor's engine, and more than half of its energy is dissipated as heat to the surroundings. The rest is used to move the tractor and dirt. The energy originally contained in a concentrated useful form in the diesel fuel is not destroyed, but it is dissipated in a dilute form, mostly to warm the surroundings very slightly.

The first law of thermodynamics must always be kept in mind in the practice of green chemistry. The best practice of green chemistry and, indeed, of all environmental science, requires the most efficient use of energy as it goes through a system. The availability of energy is often the limiting factor in using and recycling materials efficiently.

This page titled 2.5.15.2: What is Energy? is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.2: What is Energy?** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.3: Radiant Energy from the Sun

The sun is the ultimate source of most of the energy that we use. How much energy does Earth receive from the sun? If the sun were to abruptly “go out” (not to worry, it won’t happen for another billion years or so) we would quickly find out, because within hours Earth would become a frozen rock in space. In fact, the **solar flux**, which is the rate at which solar energy is transmitted through space at Earth’s distance from the sun is 1.34×103watts/m2. What this means, as illustrated in Figure 15.2, is that a 1 square meter area (a square just over 3 feet to the side) with the sun shining perpendicular to it just above Earth’s atmosphere receives energy at a rate of 1,340 watts. A watt is a measure of power, that is, energy per unit time. A power level of 1,340 watts is enough to easily power an electric iron or toaster and would provide the energy equivalent to 13100-watt incandescent bulbs plus a 40-watt bulb.

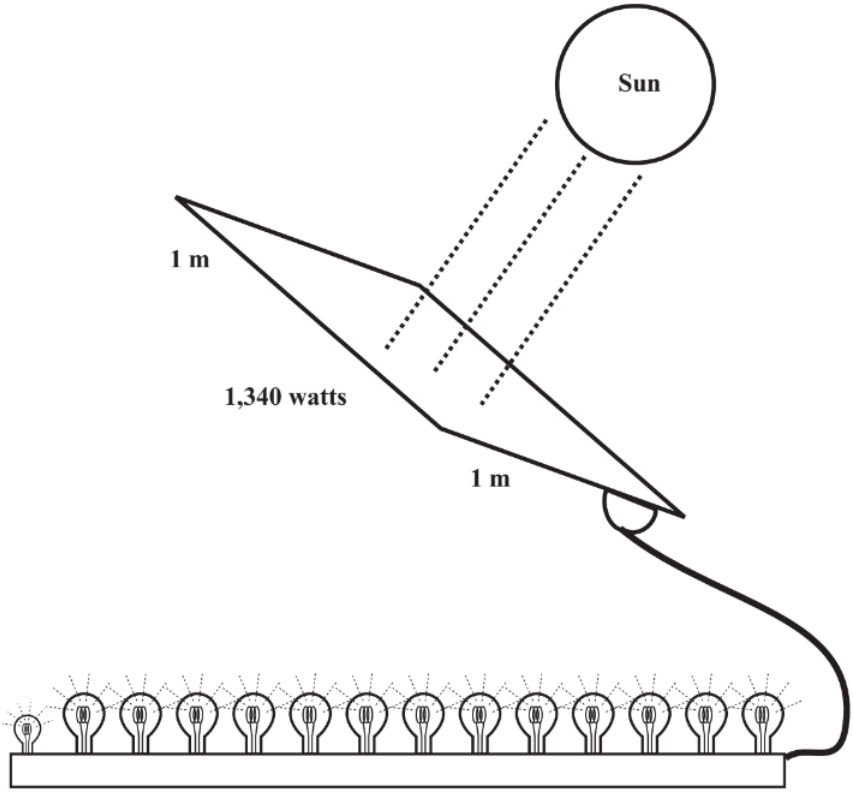


Figure 15.2. The solar flux at the distance of Earth from the sun is 1,340 watts/m2. This amount of power is equivalent to that consumed by 13 100-watt incandescent light bulbs plus a 40-watt bulb

Where does the sun get all this energy? It gets it by consuming itself in a gigantic thermonuclear fire, the same basic process that gives a “hydrogen bomb” its enormous destructive force. The fuel for the sun is ordinary hydrogen. But the energy-yielding reaction is not an ordinary chemical reaction. Instead, it is a **nuclearreaction** in which the nuclei of 4 hydrogen atoms fuse together to produce the nucleus of a helium atom of mass number 4, plus 2 positrons, subatomic particles with the same mass as the electron, but with a positive, instead of a negative, charge. There is a net loss of mass in the process (in nuclear reactions mass **can** change) and this loss translates into an enormous amount of energy. The fusion of only 1 gram of hydrogen releases as much energy as the heat from burning about 20 tons of coal. Using superscripts to express mass number and subscripts for charge, the thermonuclear fusion of hydrogen in the sun may be expressed as follows:

$$4\,{}_{1}^{1}\mathrm{H} \;\rightarrow {}_{2}^{4}\mathrm{He} + 2\,{}_{+1}^{\;0}\mathrm{e} + \text{energy}$$

This page titled 2.5.15.3: Radiant Energy from the Sun is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.3: Radiant Energy from the Sun** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.4: Sources of Energy Used in the Anthrosphere

Until about 1800 the sun provided virtually all of the energy used by humans. Biomass produced by photosynthesis was burned for heat or used to feed horses, oxen, and even human who provided muscle power. Wind resulting from uneven solar heating of air masses powered sailing ships and windmills. The solar powered hydrologic cycle produced running water that drove water wheels.

Around 1800 the dramatically increased use of coal began the era of fossil fuel energy sources. This use was enabled by the invention of the steam engine, which provided abundant and reliable power for stationary sources and locomotives and the power used for mining the coal. By 1900 coal was the dominant energy source in industrialized societies, but petroleum, especially well suited as fuel for the newly developed internal combustion engine, began a century of rapid development, becoming the favored fuel for transportation. Often encountered in wells drilled for petroleum, natural gas also developed during the 1900s, predominantly as a fuel for non-transportation needs. During this century hydroelectric power took on a significant share of electrical power production and remains a substantial renewable energy source. By around 1975, nuclear energy was supplying significant amounts of electricity and has maintained an appreciable share worldwide until the present. Miscellaneous renewable sources including geothermal and, more recently, solar and wind energy are making increasing contributions to total energy supply. Biomass still contributes significantly to the total of the sources of energy used and will become even more important as emphasis on renewable sources increases.

Figure 15.3 shows U.S. and world energy sources as of 2009. The overwhelming reliance on **fossil fuel** petroleum, natural gas, and coal are obvious. These are non-renewable energy sources. Figure 15.4 shows estimated original amounts of these fuels based on data from around 1970. Although the amounts of these fuels based upon more modern estimates would certainly differ, the proportions should be roughly the same. The predominance of coal and lignite is obvious.

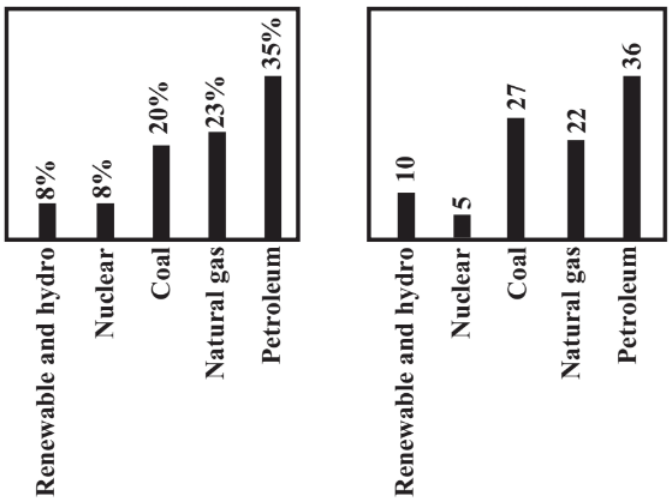


Figure 15.3. U.S. (left) and world (right) sources of energy. Percentages of total rounded to the nearest 1%.

Coal and lignite are the fuels that contribute the most carbon dioxide to the atmosphere per unit of energy generated. A measure of this contribution for fossil fuels is the ratio of water produced per molecule of carbon dioxide generated; greater relative amounts of water mean that more of the energy comes from burning chemically bound hydrogen which does not produce carbon dioxide. The best fuel in this respect is natural gas composed of methane, CH_4, which burns according to the reaction

$$CH_4 + 2\,O_2 \rightarrow CO_2 + 2\,H_2O + \text{energy}$$

producing two molecules of water per molecule of carbon dioxide. For liquid petroleum, approximate empirical formula CH_2, there is one molecule of H_2O generated per molecule of CO_2 and for coal, approximate empirical formula $CH_{0.8}$, there is somewhat less than 1/2 of a molecule of H_2O for each CO_2 molecule released.

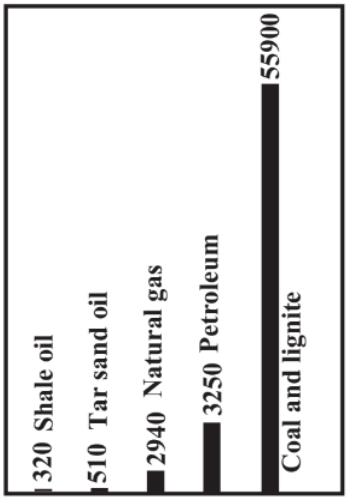


Figure 15.4. Estimated original amounts of the world's recoverable fossil fuels from 1971 (quantities in thermal petawatt hours of energy based upon data taken from M. K. Hubbert, "The Energy Resources of the Earth," in *Energy and Power*, W. H. Freeman and Co., San Francisco, 1971. A petawatt is 10^{15} watts).

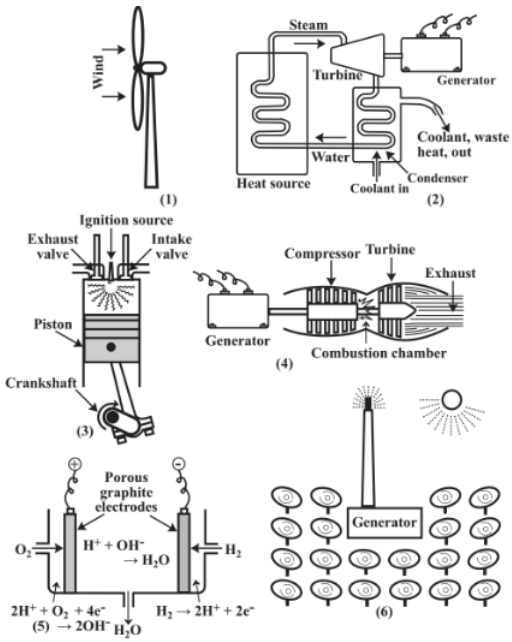
This page titled 2.5.15.4: Sources of Energy Used in the Anthrosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.4: Sources of Energy Used in the Anthrosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.5: Conversions Between Forms of Energy

The most abundant sources of energy are usually not directly useful and must be converted to other forms. Therefore, much of what is done with energy involves changing it from one form to another. As an example, the nuclear energy that can be extracted from a few kilograms of natural uranium is enormous. But in order to get any benefit from it, the uranium must first be enriched in the isotope whose nucleus can undergo fission (split) to release the energy, the enriched uranium must be placed in a nuclear reactor where fission occurs, converting the nuclear energy to heat, this heat is used to produce steam, the steam is run through a turbine to produce mechanical energy, and the turbine is coupled to a generator to convert its mechanical energy to electrical energy. The various conversions of energy from one form to another occur with different efficiencies. The successful practice of industrial ecology tries to maximize the efficiencies of energy conversion.

Figure 15.5 shows some of the most common devices for converting energy from one form to another and Figure 15.6 illustrates important energy conversions along with the percentage efficiencies with which some of these conversions can be carried out. Examination of the different percentage efficiencies for energy conversion given in Figure 15.6 shows differences ranging from very low to almost 100%. But they point to areas in which improvements may be sought. For example, photosynthesis is less than about 0.5% efficient in converting light energy to chemical energy. Despite this dismal figure, photosynthesis has generated the fossil fuels from which industrialized societies now get their energy and provides a significant fraction of energy in areas where wood and agricultural wastes are used. The intriguing possibility is suggested that genetically modified plants may be developed with much higher photochemical efficiencies, leading to greatly increased use of renewable biomass as an energy source. The poor efficiency of conversion of electricity to light in the incandescent light bulb points to the need to replace these wasteful devices with fluorescent bulbs that are 5 or 6 times more efficient.



(1) A turbine converts kinetic energy in a moving fluid to mechanical energy then to electrical energy through a generator; (2) a steam power plant in which high-energy fluid is produced by vaporizing water in a heat source such as a fossil-fueled boiler, nuclear reactor core, or solar heat collector; (3) reciprocating piston internal combustion engine; (4) internal combustion gas turbine engine from which kinetic energy of hot exhaust gases may be used to propel an aircraft; (5) fuel cell; (6) electricity generation from a solar-heated boiler and turbine generation system

Figure 15.5. Devices for conversion of energy

The most common kind of energy conversion carried out in the anthrosphere is the conversion of heat, produced by chemical combustion processes, to mechanical energy used to propel a vehicle or run an electrical generator. This occurs, for example, when gasoline in a gasoline engine burns, generating hot gases that move pistons in the engine connected to a crankshaft that converts the up-and-down movement of the pistons to rotary motion that drives a vehicle's wheels. It also occurs when hot steam generated at high pressure in a boiler flows through a turbine connected directly to an electrical generator. Unfortunately, the laws of thermodynamics dictate that the conversion of heat to mechanical energy is always much less than 100% efficient. The Carnot equation,

Extra close brace or missing open brace

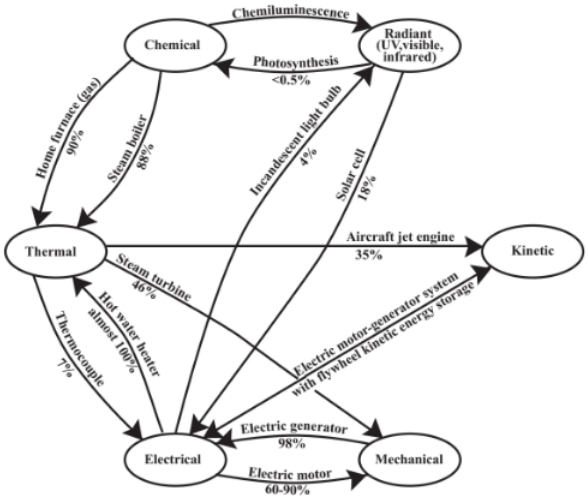


Figure 15.6. Conversion of energy between various forms and energy conversion efficiencies

states that the percent efficiency is a function of the inlet temperature (for example, of steam), T1, and the outlet temperature, T2, both expressed in Kelvin (°C + 273). Consider a steam turbine in which steam impinges on vanes attached to a rapidly rotating shaft. If the inlet temperature is 850K and the outlet temperature is 330 K, substitution into the Carnot equation gives a maximum theoretical efficiency of 61%. An inability to introduce all the steam at the highest temperature combined with friction losses of energy reduce the energy conversion efficiency of most modern steam turbines to just below 50%. Since only about 80% of the chemical energy used to raise steam by combustion of fossil fuel in a boiler is actually transferred to water to produce steam, the net efficiency for conversion of chemical energy in fossil fuels to mechanical energy to produce electricity is about 40%. Fortunately, essentially all the mechanical energy in a rotating turbine can be converted to electricity in the generator to which it is connected, so the overall efficiency of conversion of fossil fuel chemical energy to electricity is about 40%. The conversion of nuclear energy to mechanical energy in a reactor-powered steam turbine is only about 30% because reactor peak temperatures are limited for safety reason.

Another example of the application of the Carnot equation is provided by the internal combustion piston engine shown in Figure 15.5(3) in which a complete cycle consists of (1) a downward stroke sucking air into the cylinder, (2) a compression stroke during which fuel is injected, (3) ignition of the air/fuel mixture forcing the piston down, and (4) an exhaust stroke in which the exhaust gases are forced out through the open exhaust valve as the piston moves upward. The efficiency of the internal combustion engine increases with the peak temperature reached by the burning fuel, which increases with the degree of compression during the compression stroke. This temperature is highest for the diesel engine in which the compression is so high (up to around 20:1) that fuel injected into the combustion chamber ignites without a sparkplug ignition source. Whereas a standard gasoline engine is typically about 25% efficient in converting chemical energy in fuel to mechanical energy, a diesel engine is typically 37%efficient, with some reaching higher values.

Fuel Cells

Fuel cells convert the energy released by electrochemical reactions directly to electricity without going through a combustion process and electricity generator. Fuel cells are the primary means for utilizing hydrogen fuel and are becoming more common as electrical generators. The electrode reactions in a fuel cell are shown in Figure 15.5(5) and the net reaction is

$$2\,H_2 + O_2 \rightarrow 2\,H_2O + \text{electrical energy}$$

the only product of which is water. A number of different fuel cell types are at various stages of development.[1] Development is underway of solid-oxide fuel cells, operating around 1000°C that produce an exhaust that is hot enough to drive a turbine or cogenerate steam. with steam cogeneration, such systems may be able to develop overall efficiencies of up to 80%

This page titled 2.5.15.5: Conversions Between Forms of Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.5: Conversions Between Forms of Energy** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.6: Green Technology for Energy Conversion

Green technologies have much to do with the important processes by which energy is converted between various forms. Some of the more important aspects of such conversions are discussed here.

Energy Conversion **Efficiency**

Energy is best conserved by efficient energy conversion. Vastly improved energy conversion efficiencies have been achieved in heat engines such as automobile engines and gas turbines by higher combustion temperatures made possible by improved materials and heat-resistant lubricating oils. Computerized design and operation of engines enabling optimum ignition timing, valve timing, and fuel injection have made possible extremely efficient engines.

As noted in the discussion of the Carnot equation above, heat engines typically dissipate more than half of the energy in fuel as heat. A small fraction of this heat is used by heaters in automobiles. In stationary power plants much of this energy can be reclaimed for heating buildings or chemical processes with **combined power cycles**. as illustrated in Figure 15.7. Typically, in combined power cycle installations gas or fuel oil is burned in a turbine engine that is much like the engine of a turboprop airplane, and the rotating shaft of this engine is coupled to a generator to produce electricity. The hot exhaust gases from the combustion turbine can be injected into a boiler where their heat turns liquid water to steam. This steam can be run through a steam turbine coupled to a generator to produce more electricity. Steam leaving the steam turbine still contains a lot of heat, and can be conveyed to homes and other structures for heating. The water condensed from this steam is pure and is recycled to the boiler, thus minimizing the amount of makeup boiler feedwater, which requires expensive treatment to make it suitable for use in boilers. Such a system as the one described is in keeping with the best practice of industrial ecology. Heating with steam that has been through a steam turbine, a concept known as **district heating**, is commonly practiced in Europe (and many university campuses in the U.S.) and can save large amounts of fuel otherwise required for heating.

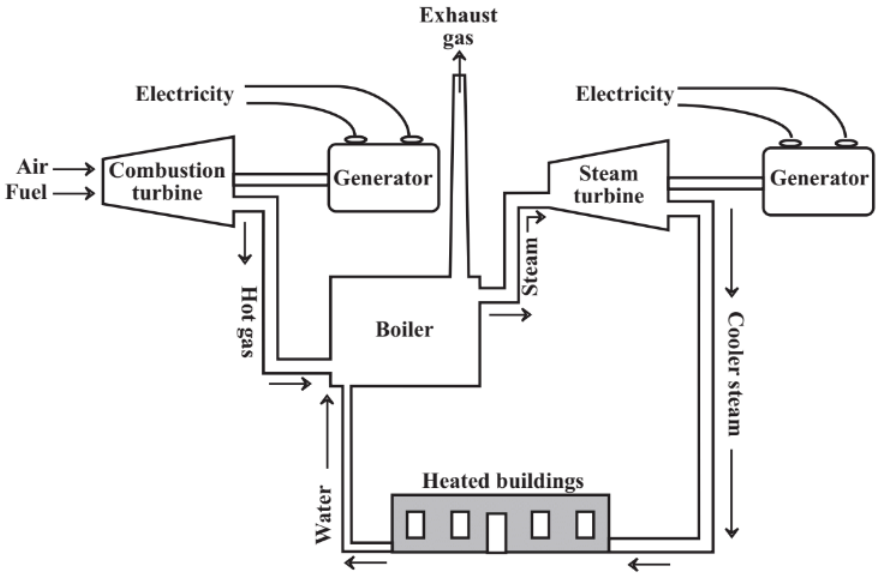


Figure 15.7. A combined power cycle in which natural gas or oil is first used to fire a gas turbine connected to an electrical generator. The hot exhaust gases from this turbine are fed to a boiler to produce steam, which drives a steam turbine, also connected to a generator. The still hot exhaust steam from the steam turbine is distributed to residential and commercial buildings for heating. The water condensed from the steam used to heat buildings is returned to the power plant to generate more steam, which conserves water and also prevents the necessity of treating more water to the high standards required for a boiler

Conversion Efficiency of Chemical Energy

In some cases a need exists to convert chemical energy from one form to another so that it can be used in a desired fashion. The generation of hydrogen gas from fossil fuels is an important chemical energy conversion process that may become much more widely practiced as fuel cells, which use elemental hydrogen as a fuel, come into more common use. Hydrogen can be obtained from a number of sources. The cheapest and most abundant raw material for hydrogen generation is coal and the same general

processes can be applied to other carbon-containing materials, especially renewable biomass. When carbon-based materials are used to generate hydrogen, the hydrogen actually comes from steam. In this process, known as **coalgasification** part of the coal is burned in an oxygen stream,

$$\mathrm{C(coal) + O_2 \ \rightarrow CO_2 + heat}$$

leaving a solid residue of very hot carbon from the unburned coal. This material reacts with water in steam,

$$\mathrm{C\,(hot) + H_2O \ \rightarrow H_2 + CO}$$

to generate elemental H_2 and CO in a reaction that absorbs heat. The CO can be reacted with more steam over an appropriate catalyst,

$$\mathrm{CO + H_2O \ \rightarrow H_2 + CO_2}$$

to increase the ratio of H_2 to CO.

The reactions shown above for the generation of elemental hydrogen from coal and water have been used for well more than a century in the coal gasification industry. Before natural gas came into common use, steam blown over heated carbon was used to generate a **synthesisgas** mixture of H_2 and CO that was piped into homes and burned for lighting and cooking. The mixture burned well, but, in addition to forming treacherous explosive mixtures with air, it was lethal to inhale because of the toxic carbon monoxide. But the process may have a future for the generation of elemental hydrogen for use in fuel cells. By using pure oxygen as an oxidant, it raises the possibility of producing greenhouse gas carbon dioxide in a concentrated form that can be pumped underground or otherwise prevented from getting into the atmosphere. Retention of carbon dioxide in this manner is called **carbonsequestration** and is the subject of some intense research.

The synthesis gas mixture of H_2 and CO_2 is a good raw material for making other chemicals, including hydrocarbons that can be used as gasoline or diesel fuel. Combined in the correct ratios over a suitable catalyst, these two gases can be used to make methane, the main constituent of natural gas:

$$\mathrm{CO + 3\,H_2 \ \rightarrow CH_4 + H_2O}$$

This page titled 2.5.15.6: Green Technology for Energy Conversion is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.6: Green Technology for Energy Conversion** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.7: Energy Conservation

Particularly with the development of an automobile-based economy and society, energy has been wasted to a shocking extent in the U.S. and some other industrialized nations. This points to the opportunity afforded by energy conservation as the most effective means of providing adequate energy. The potential for energy conservation is illustrated by the fact that some nations that have living standards near or even exceeding those of the U.S. have much lower energy consumption per capita.

On a positive note, energy use in the U.S. fell 5% to 94.6 quadrillion British thermal units (BTU) in 2009 from 99.2 quadrillion BTU in 2008, the largest year-to-year drop since the U.S.government started keeping track of energy consumption in 1949. Declines were registered in residential, commercial, industrial, and transportation sectors. Although much of the decline was due to a sharp decrease in economic activity, a significant portion was due to more energy-efficient appliances, automobiles, and other energy-consuming devices. The year also showed a significant increase in renewable wind, solar, and hydro power sources.

As illustrated in Figure 15.8, increases in the efficiency of energy utilization can contribute favorably to higher economic standards. This plot shows past and realistic future estimates of trends of the ratio of energy use per unit of gross domestic product reflecting a steady decrease in energy required per unit of economic output. Furthermore, well developed economies are more efficient users of energy. For the year 2000, the production of $1000 of gross domestic product in developed nations required 1.7 barrels of oil equivalents compared to 5.2 barrels for nations with less developed economies that lack means to utilize energy efficiently. This indicates that as the economies of less industrially advanced nations evolve with energy conservation as a top priority significant economic development with reduced energy consumption can be achieved. This kind of development combined with more efficient energy utilization in developed nations including measures such as smaller, more fuel efficient dwellings and vehicles can make a major favorable impact on energy demand.

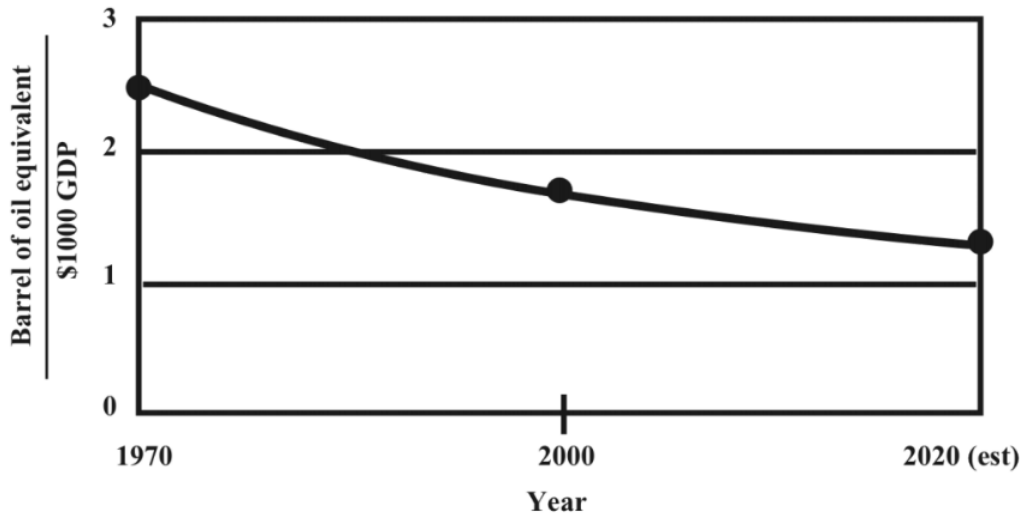


Figure 15.8. Plot of barrel of oil equivalent required per $1000 gross domestic product (GDP) as a function of year in industrialized nations.

Although household and commercial uses of energy are relatively efficient they present significant potential for increased energy efficiency. Considering the energy wasted in generating electricity, the all-electric home requires much more energy than one heated directly with fossil fuels. A compactly constructed home (more in the shape of a cube with a full basement) uses significantly less energy than a rambling “ranch style” house constructed on a slab. Apartments and row houses are much more energy-efficient than free-standing homes. Measures such as increased insulation and sealing around windows can save fuel. Waste heat from centrally located electrical power plants can be used for commercial and residential heating and cooling. With proper pollution control, these plants can use municipal refuse for a significant fraction of their fuel thereby reducing quantities of solid wastes requiring disposal.

Saving Energy in the Transportation Sector

The economic sector with the greatest potential for increased energy efficiency is transportation. Private automobiles and commercial aircraft are only about one-third as efficient as trains and buses in moving people. Movement of freight by truck

requires 5-6 times as much energy as transport by train. Furthermore, electrified railways are much better adapted to using renewable sources of energy than are private automobiles. trucks, and aircraft.

As illustrated in Figure 15.9, U.S. automobile fuel economy increased impressively from the first “energy crisis” of the 1970s until about 1990. This was achieved along with much lower emissions of exhaust pollutants. Unfortunately, the trend toward better fuel economy, which, if continued would have meant average mileage figures of at least 40 miles per gallon (MPG) by 2010, stopped moving up with increased popularity of outsized vehicles, especially the “sport utility vehicle” type. In 2007 the U.S. Congress passed legislation mandating improved fuel economy for automobiles sold in the U.S. An average of 40 MPG is readily achievable without significantly compromising safety and comfort and has the added benefit of substantially reduced emissions of greenhouse gas carbon dioxide.

Hybrid vehicles that are propelled by an electric motor connected to a battery rechargeable with a small internal combustion engine now deliver impressive fuel economy (Figure 15.10). The improvement is especially pronounced for stop-and-go driving in traffic where at least 50% better mileage figures have been achieved. In operation, the main battery in the vehicle, which is much larger than a conventional automobile battery, but smaller than one required for an all-electric vehicle, is kept charged by the onboard internal combustion engine coupled to a generator. A contribution is also made by regenerative braking that generates electricity during braking. When the vehicle is slowing, coasting downhill, or stopped, the internal combustion engine turns off, which saves fuel.

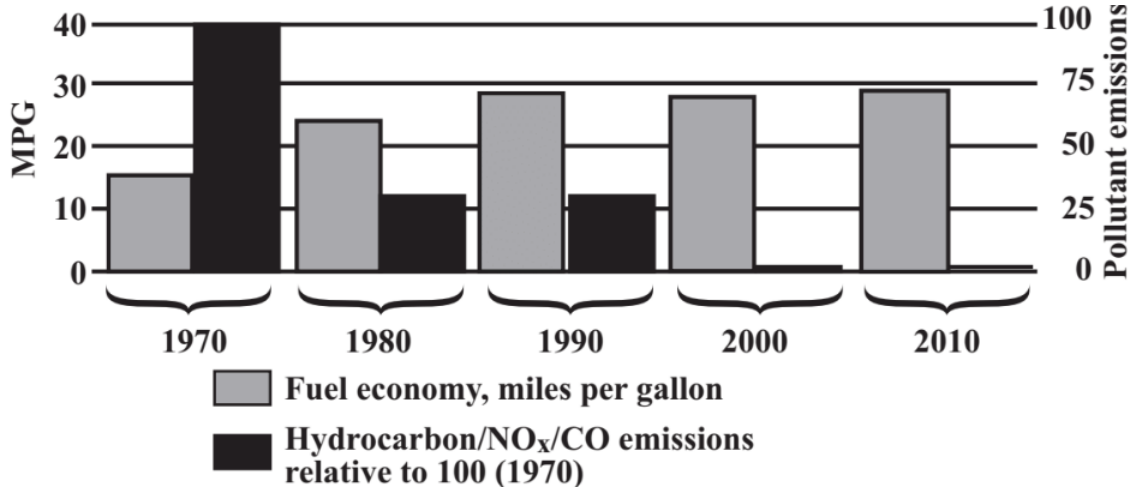


Figure 15.9. U.S. auto fleet fuel economy and emissions over 4 decades. showing a trend of generally higher fuel economy with greatly reduced emissions.

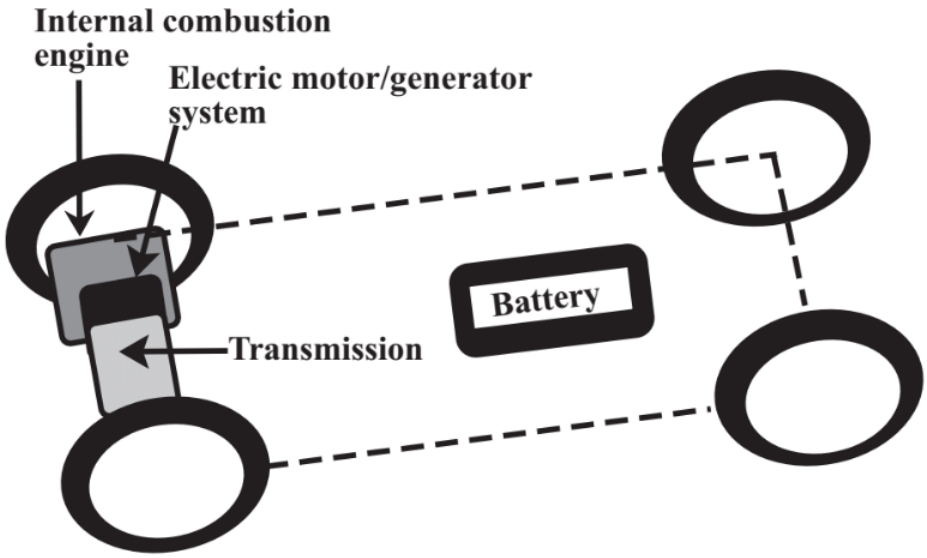


Figure 15.10. Major components of a hybrid automobile. Electrical energy to propel the car and to keep the battery charged is provided by a

generator/motor system powered by an internal combustion engine. Energy regenerated during braking is also used to help charge the battery.

Until recently the battery of choice for hybrid vehicles has been the nickel-metal-hydride battery. As of 2010, development was actively underway on hybrid vehicles using lithium ion batteries, which hold a relatively greater charge per unit mass. These batteries can be charged from an external source of electricity providing 40-50 km of driving range before the onboard internal combustion engine has to engage. Charging the battery from an external source is much less expensive than using an internal combustion engine affording rather spectacular fuel economy.

The ultimate in fuel economy could be achieved with an externally rechargeable hybrid vehicle using a diesel engine for onboard recharging. The highly efficient diesel engine, which idles on remarkably little fuel, could be left running at a steady rate for relatively long periods of time, staying hot and greatly reducing emissions, which are highest for a diesel engine during startup, shutdown, and abrupt changes in engine speed.

A natural gas (methane) fueled engine could be an excellent choice for a hybrid vehicle. With recent developments in the exploitation of abundant shale sources of natural gas, this fuel has become readily available in a number of countries including the U.S. A big advantage with methane is that it is a very low pollutant fuel. A limitation with a natural gas engine is the relatively low cruising range. But, coupled with an externally rechargeable battery in a hybrid vehicle, a natural gas engine could provide impressive range and fuel economy. The economy could be further enhanced using stratified charge ignition in which the fuel is injected directly onto the spark plug enabling extremely efficient combustion with a very lean overall fuel/air mixture.[2]

This page titled 2.5.15.7: Energy Conservation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.7: Energy Conservation** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.8: Depletable Fossil Fuels

Hydrocarbons from Wells

Currently, the most widely used sources of energy are liquid and gaseous hydrocarbons from underground. It is readily seen why liquid petroleum and natural gas became so popular because they are so easy to obtain from holes drilled into the ground, easy to transport by pipeline, and easy to use, especially in home furnaces (natural gas) and internal combustion engines (refined petroleum). But these sources suffer two major disadvantages: (1) They will eventually run out and (2) they contribute greenhouse gas carbon dioxide to the atmosphere.

Liquid petroleum occurs in pores of rock having a porosity of 10-30%, over half of which is normally occupied by water. Primary recovery typically removes about 30% of the crude oil. more advanced recovery techniques by injecting water or pressurized carbon dioxide can remove around 50% of the crude oil. A very sustainable technique is to burn petroleum in a pure oxygen atmosphere to generate electricity, collect the relatively pure CO_2 product, and inject it back into the petroleum-bearing formation to remove more oil. By using advanced techniques that recover up to 60% of the available petroleum, oil fields that have been depleted by formerly used processes can be given a second life and yield as much more oil as they did originally.

The most exciting relatively recent development in utilizing hydrocarbons pumped from below ground consists of the development of methods to remove natural gas (CH_4) from tight shale formations such as the natural-gas-rich Marcellus Shale deposits extending for 600 miles through sections of Virginia, West Virginia, Ohio, Pennsylvania, and New York. Methods have been developed for hydraulic fracturing ("hydrofracking") these formations and enabling natural gas to flow by injecting water containing additives into the formations under very high pressures. Widespread development of these sources since approximately 2000 has led to a relative abundance of natural gas. Furthermore, since about 1990 natural gas has been extracted from coal seams, many of which are not suitable for mining. There are environmental issues involving water with both of these methods of natural gas extraction. There is some potential with hydrofracking shale formations to contaminate well water with natural gas and a search of the internet can bring up rather spectacular pictures of "burning water" in which natural gas coming out of a water faucet can be lit resulting in a substantial flame. Natural gas withdrawn from coal seams is usually mixed with copious quantities of water which in some cases is polluted.

Fossil Fuels Dug from Below Ground

Large quantities of fossil fuels are dug from below ground, either from underground mines or from surface pits. The largest source of petroleum that is imported into the U.S. is from **tarsands** in the Canadian province of Alberta. These are deposits of sand covered with heavy crude petroleum that are extracted from pits and carried by enormous trucks to locations where the oil is extracted by hot water or steam leaving immense quantities of relatively clean sand.

Another petroleum substitute that can be obtained by mining is **shale oil**, a material bound to oil shale rock in the form of complex organic material of biological origin called kerogen. Shale oil is removed from these rocks by heating in the absence of air. It is believed that as much as1.8.trillion barrels of shale oil could be recovered from deposits in Colorado, Wyoming, and Utah. The story is told that oil shale was discovered in this region by an earlier settler who constructed a fireplace and chimney in a cabin from the shale and was very much distressed when the first fire in the heating facility resulted in its burning, along with the surrounding cabin. Although large amounts of petroleum substitute could be recovered from oil shale, this resource is not likely to be developed on a large scale because the pyrolysis releases enormous amounts of greenhouse gas carbon dioxide and leaves a residue of ash from which salts such as sodium sulfate are readily leached. Furthermore, the liquid shale oil product contains a high content of potentially carcinogenic organonitrogen compounds.

Coal and Lignite

Coal and related solids are solid carbonaceous fossil fuels formed by the partial biodegradation of ancient biomass followed by geochemical processes involving heat and high pressures. Coal is differentiated largely by **coalrank** based upon percentage of fixed carbon, percentage of volatile matter, and heating value. Although "coal" gives the appearance of being pure carbon, it is actually a complex hydrocarbon-like material, typically with an empirical formula of around $CH_{0.8}$ and containing from 1 to several percent sulfur, nitrogen, and oxygen. Of these elements, sulfur bound to the organic coal molecule and mixed with coal as mineral pyrite, FeS_2, presents major environmental problems because of production of air pollutant sulfur dioxide during

combustion. Much of the FeS_2 can be removed physically from coal prior to combustion and sulfur dioxide can be removed from stack gas by various scrubbing processes. Lower rank brown coal and lignite typically have high moisture and bound oxygen contents. Most commonly, coal classified as brown coal is relatively closer in its constitution to the vegetation from which it was formed.

The greatest fraction of electricity production worldwide is from coal burned in boilers to raise steam that runs turbines connected to electrical generators. The potential is high for air pollution from this technique including especially fly ash and sulfur dioxide. These two pollutants are now generally well controlled. However, burning coal releases more greenhouse gas carbon dioxide per unit energy output than any other energy-yielding process and this release can only be controlled by extraordinary (and expensive) means.

Coal Conversion

A sustainable approach to utilizing the coal energy resource is **coalconversion** in which coal is converted to gaseous or liquid fuels or low-sulfur solids. Starting with the house of William Murdocks at Redruth, Cornwall, England, illuminated with coal gas in 1792, coal conversion has along history. Pall Mall in London was lit with gas from the first municipal coal-gas system in1807. The coal-gas industry began in the U.S. in 1816. The first coal gas plants operated by heating coal in the absence of air leaving a solid residue consisting mostly of carbon (that could be used as stove fuel) and produced a hydrocarbon-rich fuel especially effective for lighting. During the 1800s a gasification process was developed in which steam reacted with hot carbon to produce a mixture consisting primarily of H_2 and CO (synthesis gas) to which it was necessary to add volatile hydrocarbons to make the fuel suitable for lighting. There were 11,000 coal gasifiers operating in the U.S. in the 1920s and the industry peaked in 1947 after which it was rapidly replaced by abundant natural gas sources. During World War II Germany made synthetic petroleum from synthesis gas, reaching a capacity of 100,000 barrels per day in 1944. The largest coal-based synthetic fuels plant operating today is in Sasol, South Africa, and now produces hydrocarbons and feedstocks equivalent to about 150,000 barrels of petroleum per day.

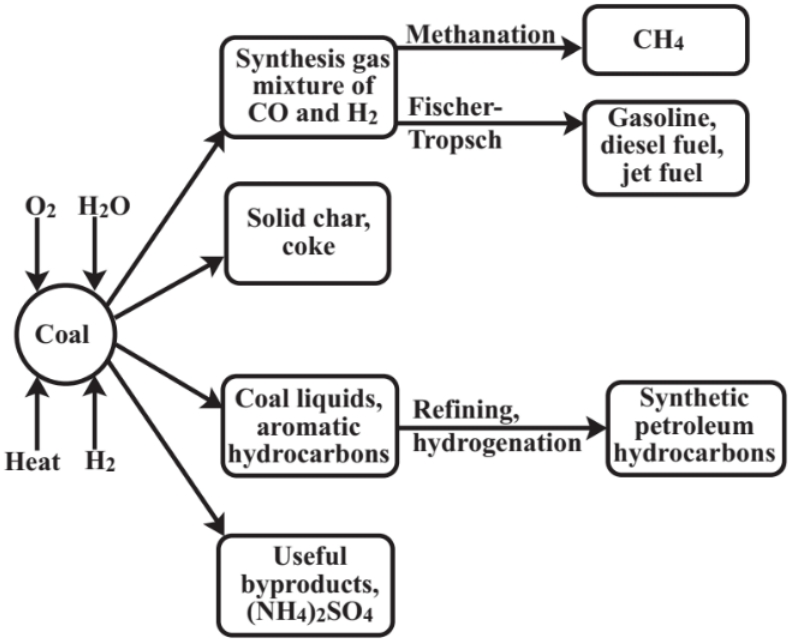


Figure 15.11. Routes to coal conversion. BTU refers to British thermal units, a measure of the heat energy that can be obtained from a fuel. Methanation means synthesis of CH_4 gas. Hydrogenation and hydrotreating refer to reaction with elemental H_2 gas.

Although coal conversion could be developed as a substitute for petroleum, it is by no means a green process considering the environmental costs of coal mining, the production of toxic coal tar byproducts, and the enormous amounts of carbon dioxide generated during the conversion process. As discussed in the following section, the byproduct carbon dioxide can be captured and pumped underground where it can aid in petroleum recover.

This page titled 2.5.15.8: Depletable Fossil Fuels is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.8: Depletable Fossil Fuels** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.9: Carbon Sequestration for Fossil Fuel Utilization

Fossil fuels can be used sustainably by employing **carbonsequestration** to retain greenhouse gas carbon dioxide produced in fuel combustion or coal conversion, which by preventing carbon dioxide generated by fossil fuels from entering the atmosphere holds the promise of enabling utilization of fossil fuels without contributing to greenhouse warming. Carbon dioxide can be captured and sequestered deep in the ocean or pumped into porous formations underground. A major concern with oceanic disposal of carbon dioxide is the tendency of this gas to lower pH resulting in thinner shells in shellfish.

The most promising approach to carbon sequestration is to pump carbon dioxide gas into underground formations at depths exceeding 1000 meters. If the formations are overlain with impermeable layers of rock that are not breached by improperly abandoned oil wells, the carbon dioxide will remain in place indefinitely. Sequestration is aided by the presence of saline groundwater and by chemical reaction with minerals.

Carbon dioxide from natural gas that contains a high content of CO_2 has been sequestered since 1996 in the Sleipner oil and gas field about 240 km off the Norwegian coast. The gas is pumped into the 200-m thick Utsira sandstone formation located about 1000 km below the seabed. A mixture of carbon dioxide and toxic hydrogen sulfide is now being disposed underground in Alberta, Canada.

Carbon dioxide is best captured and sequestered in processes such as fermentation of sugars to make ethanol that produce the gas in high concentrations. Because of the high nitrogen contents exhaust gases from combustion are not suitable for carbon dioxide sequestration. Using pure oxygen instead of air for combustion does produce a relatively pure carbon dioxide product and one commercial power plant using coal burned in pure oxygen is now planned with support of the U.S. Department of Energy.

The most promising approach for large-scale carbon dioxide sequestration is through coal gasification (see Section 15.8). There are two major sources of carbon dioxide from coal gasification. The first of these is coal combustion with pure oxygen oxidant,

$$\mathrm{C(coal) + O_2 \ \rightarrow CO_2 + heat}$$

which generates the heat required for gasification of the hot carbon residue of coal with steam:

$$\mathrm{C(coal) + H_2O \ \rightarrow CO + H_2}$$

The second reaction that produces carbon dioxide is reaction of steam with CO to increase the ratio of H_2 to CO in the synthesis gas product:

$$\mathrm{CO + H_2O \ \rightarrow H_2 + CO_2}$$

The largest carbon dioxide sequestration process now operating in the U.S. is the Great Plains Synfuels Plant near Beulah, North Dakota. This plant gasifies 16,000 tons per day of lignite coal and sends approximately 4.3 million cubic meters of carbon dioxide per day (3 million tons per year) through a 330 km pipeline to the Weyburn and Midale oilfields in Saskatchewan, Canada, for sequestration and petroleum recovery.

An integrated coal gasification plant with carbon dioxide sequestration is shown in Figure 15.12. This plant uses the reaction of steam with hot carbon from coal to produce elemental hydrogen (reaction 15.9.2) and reacts the CO product with steam to produce more H_2; both of

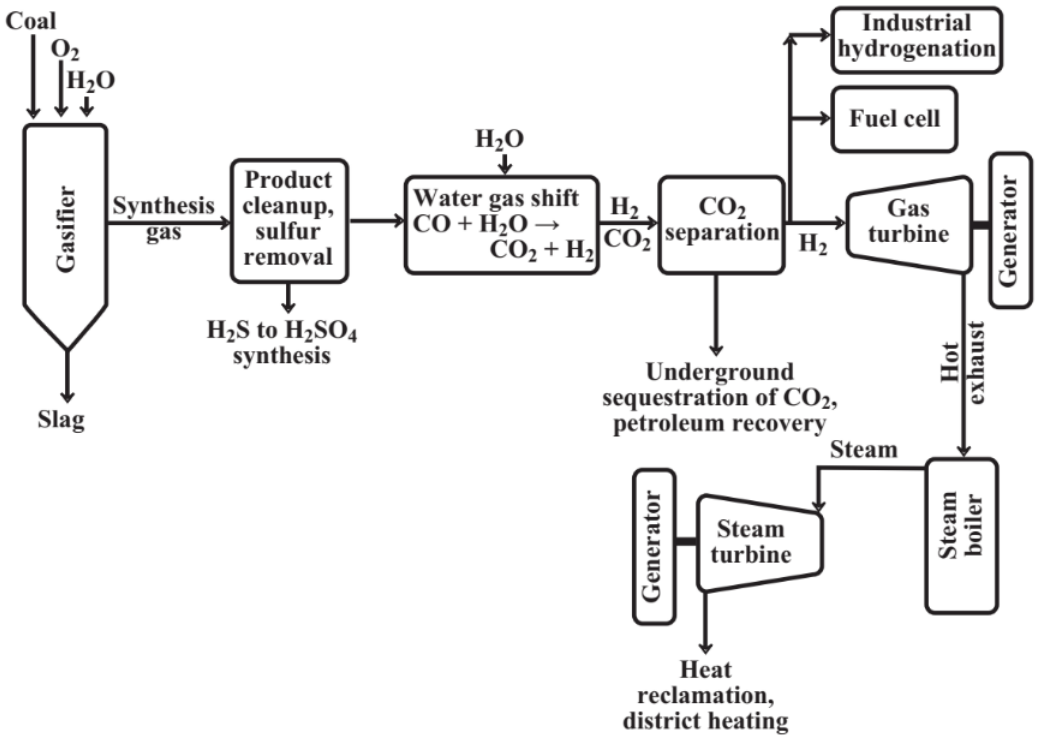


Figure 15.12. An integrated coal gasification plant with carbon dioxide sequestration

these reactions produce CO_2. A gas turbine fueled with H_2 coupled to a generator produces electricity. Hot exhaust gas from the gas turbine is used to raise steam in a boiler and this steam powers a steam turbine that is also coupled to a generator. This combination results in very efficient electrical power generation. The elemental H_2 generated can be used in fuel cells, as a chemical reagent to produce synthetic hydrocarbon fuels, or to synthesize ammonia, NH_3. Some ammonia is also produced from nitrogen in the coal during gasification and is recovered as a product. Sulfur, which occurs in essentially all coals, is released during gasification as byproduct hydrogen sulfide used to make sulfuric acid or disposed with the carbon dioxide byproduct. The CO_2 is separated from the gas and pumped under high pressure into mineral formations at depths up to around 2000 m. If these formations are oil-bearing, the carbon dioxide enables petroleum recovery. Byproduct heat from the plant can be used for district heating. Particularly when it is integrated with byproduct recovery, chemical synthesis of NH_3 and H_2SO_4, and district heating, an integrated plant of this kind is an excellent example of a system of industrial ecology.

Fuel from Carbon Dioxide

An interesting potential use of carbon dioxide is to use it as a source of carbon for the synthesis of hydrocarbon fuels and other organic compounds including alcohols. The requirement for so doing is an abundant and inexpensive source of elemental hydrogen, H_2, which can be reacted with CO_2 through the reverse water-gas shift reaction:[3]

$$CO_2 + H_2 \rightarrow CO + H_2O$$

The CO produced can be reacted with additional H_2 to produce methane gas (methanation), hydrocarbon liquids (Fischer-Tropsch synthesis), or alcohols. The overall process is the reverse of the energy-yielding combustion of hydrocarbon fuels, so it consumes a lot of energy and requires a cheap source of energy in order to be practical. The most likely energy source is “free” wind power generating electricity that can produce H_2 gas by electrolysis of water (see Reaction 15.11.1). In this application the intermittent characteristic of wind power is not an issue and hydrogen produced in abundance during times of strong wind can be pumped underground to be withdrawn for subsequent reaction with CO_2 and CO.

This page titled 2.5.15.9: Carbon Sequestration for Fossil Fuel Utilization is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.9: Carbon Sequestration for Fossil Fuel Utilization** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.10: Nuclear Energy

Nuclear energy is generated by the neutron-induced fissioning of heavy atomic nuclei, most commonly those of the uranium isotope with a mass number of 235 or plutonium with a mass number of 239, to produce radioactive fission products, an average of 2.5 more neutrons and an astounding amount of energy compared to an ordinary chemical reaction. A typical example of such a fission reaction is the following:

$$^{235}_{92}\mathrm{U} + {}^{1}_{0}\mathrm{n} \rightarrow {}^{133}_{51}\mathrm{Sb} + {}^{99}_{41}\mathrm{Nb} + 4\,{}^{1}_{0}\mathrm{n}$$

A nuclear reactor operating at a constant power level is controlled such that on average 1 neutron from each fission reaction is absorbed to cause another fission reaction, thus sustaining a **chain reaction**. The excess neutrons are absorbed by nonfissionable material. In order to cause the desired fission, the neutrons, initially released as rapidly moving, high energy particles must be slowed down, which is done by a **moderator**, such as water, in the reactor.

The basic function of a nuclear power reactor is to serve as a heat source to produce steam used to generate mechanical energy. The basic components of a nuclear power reactor are shown in Figure 15.13. Pressurized superheated water circulates through the hot reactor core in an enclosed loop (to prevent escape of radioactive contaminants). Heat from this water is used to convert water to steam in a heat exchanger. The rest of the power plant is like a conventional fossil-fueled plant with a steam turbine coupled to a generator and the steam from the steam turbine being condensed to provide liquid water for the heat exchanger.

Although only 0.71% of natural uranium is fissionable uranium-235, and uranium to be used for fission must be enriched in this isotope, there is an adequate global supply of uranium. In principle, the remaining 99.28% of uranium that consists of uranium-238 could be converted to fissionable plutonium by absorption of neutrons in **breeder reactors**. Plutonium is actually generated by uranium-238 absorbing neutrons in a conventional nuclear power reactor, and after the reactor has operated for a few months after refueling, a large fraction of its energy output comes from plutonium generated in the reactor.

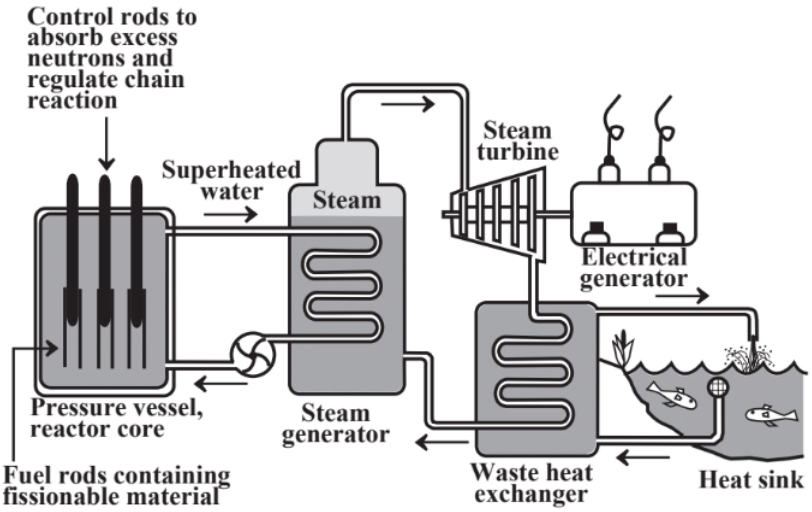


Figure 15.13. Nuclear fission reactor for electrical power generation.

An interesting possibility for breeder reactors is the liquid-sodium-cooled **travellingwave reactor** in which a relatively small segment at the end of a mass of uranium held in an array of fuel rods is enriched in fissionable uranium-235. The fission process is initiated in this portion of the fuel and neutrons from it migrate to the adjacent non-enriched segment where they are absorbed by uranium-238 to produce fissionable plutonium-239. This builds up a sector containing enough plutonium to sustain fission and the nuclear fission process very slowly and continuously migrates into the area with the newly fissionable material. With proper design such a reactor can function for decades before the “wave” reaches the end of the mass of uranium and no more fissionable material is produced to sustain fission.

A major issue with nuclear power reactors is posed by the **high-level nuclear wastes** consisting of radioactive fission products generated when the uranium nucleus splits apart and the radioactive transuranic elements produced when uranium nuclei absorb neutrons. Some of the radioactive fission products may last for several centuries and some of the transuranic isotopes remain lethal

for thousands of years. At the present time, spent fuel elements are being stored under water at the reactor sites. This is actually a good thing because the short-lived fission products that are responsible for most of the radioactivity in nuclear fuel freshly removed from a reactor decay rapidly, and after a few years of storage only a small fraction of the original activity is present. Under current regulations in most countries, the wastes from this fuel will eventually have to be buried. A better alternative is to process the material in the spent fuel elements to remove radioactive products from uranium fuel. The relatively short-lived fission products decay spontaneously within several hundred years and can be stored in a secure location for several centuries. The longer-lived nuclear transuranic wastes can be bombarded with neutrons in nuclear reactors, a process of **transmutation** in which the elements are converted to other elements or fission products with shorter half lives resulting in relatively rapid production of stable isotopes. Radioactive waste elements for which transmutation is feasible include plutonium, americium, neptunium, curium, technetium-99, and iodine-129. Plutonium, americium, neptunium, and curium are heavy actinide elements that are fissionable and add fuel value in a nuclear reactor.

Another problem with nuclear reactors is their decommissioning. One option is to dismantle the reactor soon after it is shut down using apparatus operated by remote control. The radioactive reactor parts are then disposed. Another approach is to allow the reactor to stand for 30–100 years before dismantling, by which time most of the radioactivity has decayed (and most of the people responsible for the reactor initially have died). A third option is to entomb the reactor in a concrete structure.

Two accidents have dealt a strong blow to the future of nuclear energy. The first, and much lesser of these, occurred on March 28, 1979, when Metropolitan Edison Company's nuclear reactor located on Three Mile Island in the Susquehanna River, 28 miles outside of Harrisburg, Pennsylvania, lost much of its coolant resulting in overheating, and partial disintegration of the reactor core. Some radioactive xenon and krypton gases were released to the atmosphere and some radioactive water entered the river. The problem was remediated and the reactor building sealed. Then in April of 1986 a reactor of inherently dangerous Soviet design blew up in Chernobyl, which is now part of Ukraine. Officially, 31 people were killed, but the death toll was probably many more, especially when delayed effects of exposure to radioactive materials are considered. Food, including reindeer meat in Lapland, was contaminated as far away as Scandinavia, thousands of people were evacuated, and the entire reactor building was entombed in a massive concrete structure. The reactor that blew up was one of four units, the last of which was not shut down permanently until the end of 2000!

Given the horrors described above, why would reputable scientists even advocate development of nuclear energy? The answer is, simply, carbon dioxide. With massive world resources of coal and other non petroleum fossil fuels, the world has at least enough readily available fossil fuel to last for a century. But evidence is mounting that the carbon dioxide from fossil fuel combustion is leading to global warming accompanied by effects such as rising sea levels that will inundate many coastal cities. Humans do know how to design and operate nuclear reactors safely and reliably; indeed, France has done so for years and gets most of its electricity from nuclear fission, and the U.S. Navy has had an exemplary safety record with reactors on submarines and aircraft carriers. So, it may be that nuclear energy is far from dead and that humankind, reluctantly and with great care, will have to rely on it as the major source of energy in the future. A new generation of nuclear power plants is waiting to be built that have the desirable characteristics of **passive stability**. This means that measures such as gravity feeding of coolant, evaporation of water, or convection flow of fluids operating automatically provide for safe operation of the reactor and automatic shutdown of the reactor if something goes wrong. New designs are also much more reliable with only about half as many pumps, pipes, and heat exchangers as are contained in older power reactors.

Nuclear Fusion

The fusion of a deuterium nucleus and a tritium nucleus releases a lot of energy as shown below, where Mev stands for million electron volts, a unit of energy:

$$ {}^{2}_{1}\mathrm{H} + {}^{3}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He} + {}^{1}_{0}\mathrm{n} + 17.6\,\mathrm{MeV}\,(\text{energy released per fusion}) $$

This reaction is responsible for the enormous explosive power of the "hydrogen bomb." So far it has eluded efforts at containment for a practical continuous source of energy. And since physicists have been trying to make it work on a practical basis for the last approximately 60 years, it will probably never be done. (Within about 15 years after the discovery of the phenomenon of nuclear fission, it was being used in a power reactor to propel a nuclear submarine.) However, the tantalizing possibility of using the essentially limitless supply of deuterium, an isotope of hydrogen, from Earth's oceans for nuclear fusion still give some investigators hope of a practical nuclear fusion reactor.

Nuclear fusion was the subject of one of the greatest scientific embarrassments of modern times when investigators at the University of Utah in 1989 announced that they had accomplished so-called cold fusion of deuterium during the electrolysis of deuterium oxide (heavy water). The announced “discovery” of cold fusion resulted in an astonishing flurry of activity as scientists throughout the world sought to repeat the results, whereas others ridiculed the idea. Unfortunately, for the attainment of a cheap and abundant source of energy, the skeptics were right, and the whole story of cold fusion stands as a lesson in the (temporary) triumph of wishful technological thinking over scientific good sense.

This page titled 2.5.15.10: Nuclear Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.10: Nuclear Energy** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.11: Renewable Energy Sources - Solar Energy

Ideal energy sources are those that do not pollute and never run out. Such sources are commonly called renewable energy resources. There are several practical renewable energy resources that are discussed briefly in this section. The sun is the ultimate renewable energy source and solar energy is discussed first.

Solar Energy is The Best — When The Sun Shines

Sunshine comes close to meeting the criteria of an ideal energy source, including widespread availability, an unlimited supply, and zero cost up to the point of collection. The utilization of solar energy does not cause air, heat, or water pollution. Sunshine is intense and widely available in many parts of the world. If it were possible to collect solar energy with a collection efficiency of 10%, approximately one-tenth of the area of Arizona would suffice to meet U.S. energy needs, and at 30% collection efficiency, only about one-thirtieth of the area of that generally sunny state would suffice. But, keep in mind that such an area is still enormous and the implications of covering it with solar collectors would be profound.

There are several ways in which solar energy can be utilized. The simplest of these is for heating, and solar-heated houses and solar water heaters have been developed and used successfully. At a somewhat more sophisticated level, solar boilers have been developed that are located on towers and receive concentrated sunlight from an array of parabolic mirrors, thus generating steam to make electricity (see Figure 15.5). Some years ago a serious proposal was even made to use solar collectors in Earth orbit and convert the energy to a beam of microwave radiation focused on a receiver on Earth’s surface. Visions of this beam straying from its aiming point or hapless birds or even aircraft straying into it and being instantly cooked by an extraterrestrial microwave oven have prevented this plan from coming to fruition. Photosynthetic generation of biomass is another way of utilizing solar energy as discussed in a later section of this chapter.

Other than low-grade building and water heating, the most promising way to utilize solar energy is by its direct conversion to electricity in photovoltaic cells (see Figure 15.14). Originally just a laboratory curiosity, these devices became practical sources of electricity for satellites and space vehicles where their high cost was of little concern. But over the years they have become more efficient and cheaper, and it is now common to see arrays of these cells used to power data processors and signaling devices in remote locations. And some houses even have banks of photovoltaic cells.

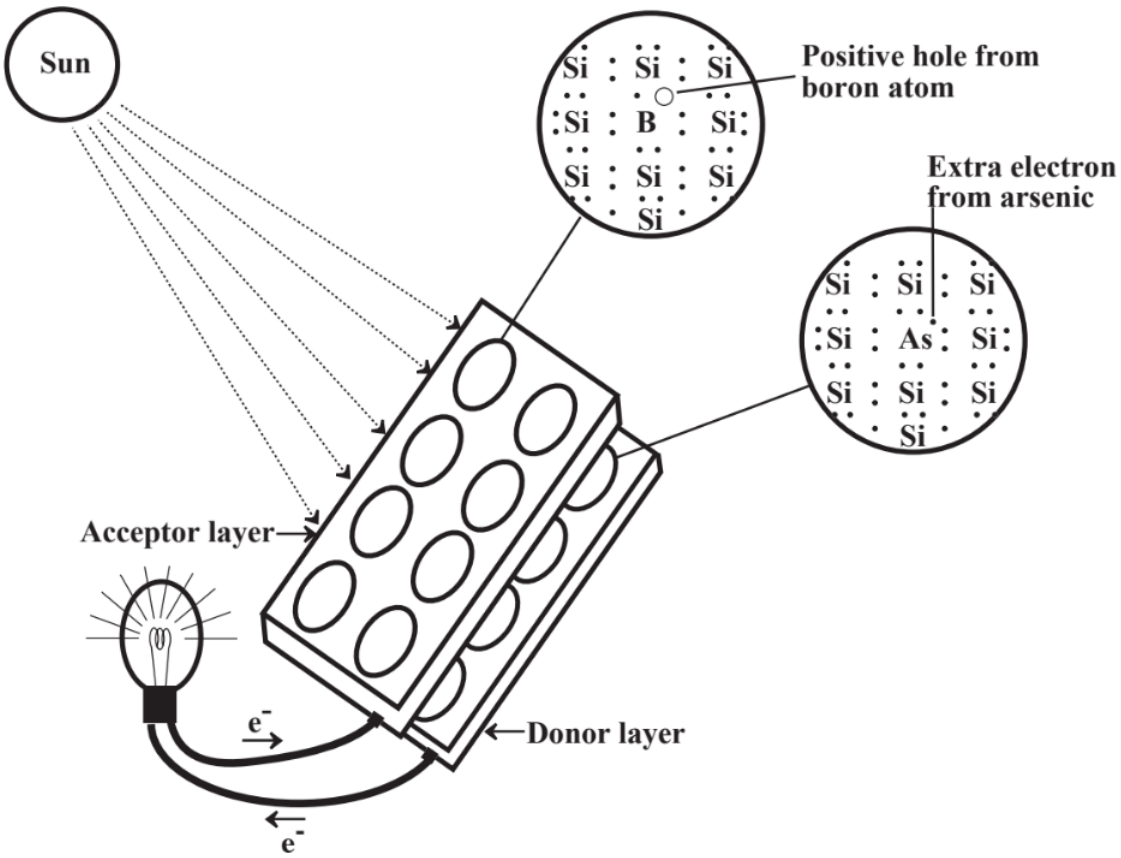
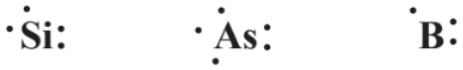


Figure 15.14.

Photovoltaic cell

Photovoltaic cells depend upon the special electronic properties of silicon atoms containing low levels of other elements. The cell consists of two layers of silicon, a donor layer that is doped with about 1 part per million of arsenic atoms and an acceptor layer doped with about 1 part per million of boron. Examination of the Lewis symbols of these three elements,



shows that substitution of an arsenic atom with its 5 valence electrons for a silicon atom with its 4 valence electrons in the donor layer gives a site with an excess of 1 electron whereas substitution of a boron atom with only 3 electrons for a silicon atom in the acceptor layer gives a site “hole” that is deficient in one electron. The surface of a donor layer in contact with an acceptor layer contains electrons that are attracted to the acceptor layer. When light shines on this area, the energy of the photons of light can push these electrons back onto the donor layer, from which they can go through an external circuit back to the acceptor layer. This flow of electrons constitutes an electrical current that can be used for energy.

Current photovoltaic cells are around 12–15% efficient in converting radiant solar energy to electricity at a cost significantly higher than that of electricity generated in fossil fuel powerplants. However, advances are continually being made in solar cell technology and it can be anticipated that efficiencies will continue to increase as costs decrease. The obvious major disadvantage of solar energy is that it does not work in darkness, and variable atmospheric conditions affect its output. Flexibility in electrical power grids allows such intermittent sources for up to 15% of power without using special devices for energy storage. Furthermore, there are means of storing energy, such as by extremely high-temperature/high-pressure supercritical water stored deep underground or mechanical energy stored in the extremely rapid rotation of flywheels.

A very attractive energy storage option for solar energy given the growing use of fuel cells is hydrogen gas. Electrolysis of water containing a solution of electrolyte (commonly KOH)

$$2\,H_2O + \text{electrical energy} \rightarrow 2\,H_2\,(g) + O_2\,(g)$$

with solar-generated electricity provides elemental hydrogen and oxygen, which are exactly the fuels used by fuel cells. Commercially available electrolyzers are 55-75% efficient in converting electricity to hydrogen and oxygen. The overall efficiency of this process can be increased significantly by the development of direct means for splitting water molecules into hydrogen and oxygen using the energy of light photons.

This page titled 2.5.15.11: Renewable Energy Sources - Solar Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.11: Renewable Energy Sources - Solar Energy** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.12: Energy from Wind and Water

Moving fluids are rich sources of energy that can be tapped, usually by turbines linked to electrical generators. Moving water and blowing winds are energy-rich fluids that have some similarities as energy sources so they are considered together in this section. Both are renewable and both are indirect means of harnessing solar energy — winds produced by the uneven heating of air masses and water carried by the solar-powered hydrologic cycle. Both are among the oldest sources of energy, such as wind used to propel sailing ships and waterwheels used for centuries to grind grain. And both are among the newest sources of energy — winds with technologically advanced wind turbines and water through ingenious devices such as those used to capture the energy of moving water in ocean tides.

Favorable Winds

As the sun heats air masses unevenly, winds are generated that can be tapped as an indirect form of solar energy. Wind power is undergoing rapid growth in a number of countries and has become competitive in cost with more conventional sources in some areas. In parts of Europe, California, Wyoming, and other locations, the sight of wind-powered generators mounted on towers has become common (Figure 15.15). In 2009 world wind power capacity increased by 31%reaching a capacity of 158 gigawatts. During 2009 China's wind power capacity doubled from 12GW to 25 GW and the U.S. capacity grew by 10 GW to 35 GW total capacity. This gave the U.S.the largest wind power capacity of any nation in the world, though rapidly being overtaken by China.

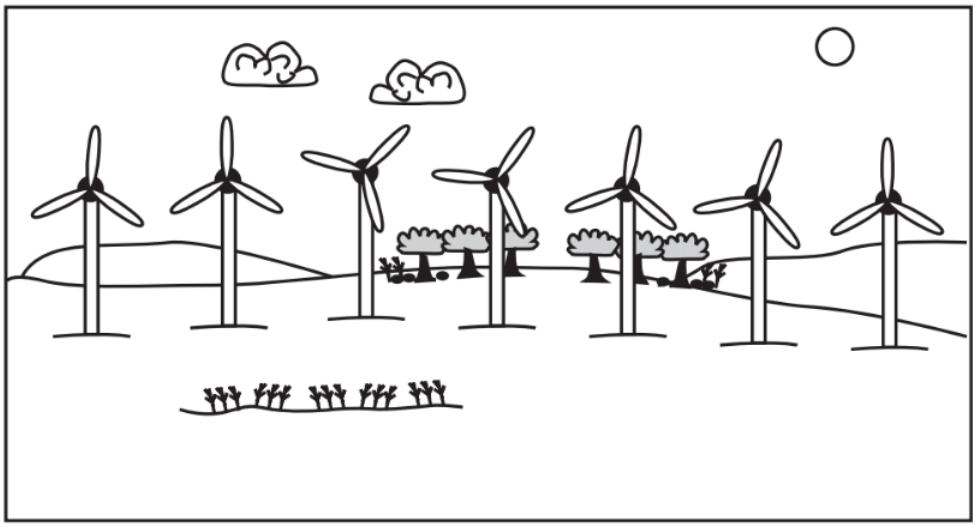


Figure 15.15. Wind-powered electrical generators mounted on towers are becoming increasingly common sights in the world in areas where consistent wind makes this nonpolluting source of renewable energy practical.

Modern wind turbines are generally large and sophisticated machines with diameters of 40 to 50 meters and rated between 0.5 to 2 megawatts. The largest turbines available as of 2010 were rated at 7 MW. Wind turbines are designed to operate consistently at varying wind speeds, to respond to changing wind direction, and to operate over a wide range of temperatures. Provision is made for electrical resistance heating of turbine blades in cold climates where ice accumulation is likely.

Offshore locations with turbines firmly anchored to the sea bed offer several advantages for production of wind power including generally consistent winds in coastal areas and lack of conflict with uses of land. The largest offshore wind power project to date began operation off the coast of southeast England in September 2010. This facility operated by Vattenfall, a Swedish energy company, has 100 turbines each 115 meters tall, and produces up to 300 megawatts of electricity, enough to power 200,000 homes

Energy from Moving Water

Various means of harnessing the energy of moving water have been used since ancient times with water-powered sailing vessels and waterwheels driven by flowing water for grinding grain known in ancient Greece and Rome. Waterwheels up to 50 horsepower were developed in the Middle Ages and were widely applied to grist mills, sawmills, and leather, textile, and machine shop

operations in Colonial America. These sources suffered from problems due to irregular water flow and freezing during winter so that many of these facilities were abandoned when steam engines became widely available in the early 1800s.

A renaissance in waterpower occurred in the late 1800s when it became practical to drive electrical generators with water. Starting with the first hydroelectric plant in the U.S. on the Fox River near Appleton, Wisconsin, in 1882, hydroelectric power driven by sophisticated power turbines grew rapidly in the U.S. and throughout the world. By 1980 hydroelectric power composed 25% of world electricity production and 5% of total world energy generation. Norway generates essentially all of its electric power and about 50% of its total energy from hydroelectric sources.

The morphology of the geosphere has a strong influence on the potential for hydroelectric power generation, which is favored by mountainous terrain and large river valleys. Abundant rainfall and snowmelt are also important factors. China has about 1/10 of the world's potential for hydroelectric energy and its enormous Tree Gorges installation on the Yangtze River is the world's largest. The reservoir for this power source has been impounded with a 185-meter high, 1,300-m long dam located at the end of a number of steep canyons holding a body of water that extends for 630 km with an average width of 1.2 km. When fully operational, this massive installation will have 32 generating units and a capacity of 22.5 gigawatts, equivalent to 22 large coal-fired powerplants.

With the hydrologic cycle continuously pumping water into it, hydroelectric power is certainly sustainable, and prevents release of greenhouse gases. Bodies of water impounded to provide power can serve as water supplies for agriculture, municipalities, and industries. The potential exists to practice aquaculture in reservoirs by raising fish and freshwater shrimp(prawns). On the negative side, the development of hydroelectric power can present some serious environmental problems. In the modern era construction of a large power reservoir displaces significant numbers of people (more than 1 million for China's Three Gorges project), alters river flow, changes aquatic ecology, and fills once scenic valleys with water. In several significant cases dams have been removed from rivers to restore their valleys to their former state.

Water Energy without Dams

Hydrokinetic and **wave energy conversion** devices are being developed to harvest the kinetic energy of moving water in natural streams, tidal estuaries, ocean currents, and constructed waterways free of dams. A typical such device consists of a turbine with relatively large and widely spaced blades coupled directly to a generator that is fixed in a river or other water current. Such a device can be anchored directly to a river bed or attached to bridge supports.

Another source of energy from moving water is that from tides, changing levels of seawater resulting from the gravitational pull of the sun and moon. Tidal energy is feasible as demonstrated by the 240 megawatt tidal power station that has operated reliably in the Rance estuary region of France since it was constructed in 1966. This facility has about 1/4 the capacity of a standard 1,000MW coal-fired or nuclear plant. Several other small installations have been built including an 18MW experimental unit at Annapolis Royal, Canada. Tidal electricity generating stations suffer from the disadvantage that sufficient water flows to generate electricity only about 10 hours per day. Nevertheless, the amount of energy potentially available from tides is enormous and it is completely renewable.

An interesting way of harnessing water energy is **pressure-retardedosmosis** in which saline ocean water and fresh water are separated by a water-permeable membrane and the flow of water through the membrane from the fresh water to the saline water side builds pressure in the latter that can be harnessed to produce electricity. Pressure-retarded osmosis is illustrated in Figure15.16. Although the process operates on a continuous basis it is shown as a stepwise process in Figure 15.16 to illustrate the operating principle. The world's first osmotic plant, a demonstration unit with a minuscule capacity, went into operation in Tofte, Norway, in November, 2009. Pressure-retarded osmosis plants can be located in almost any of the huge number of locations worldwide where fresh water flows into the sea.

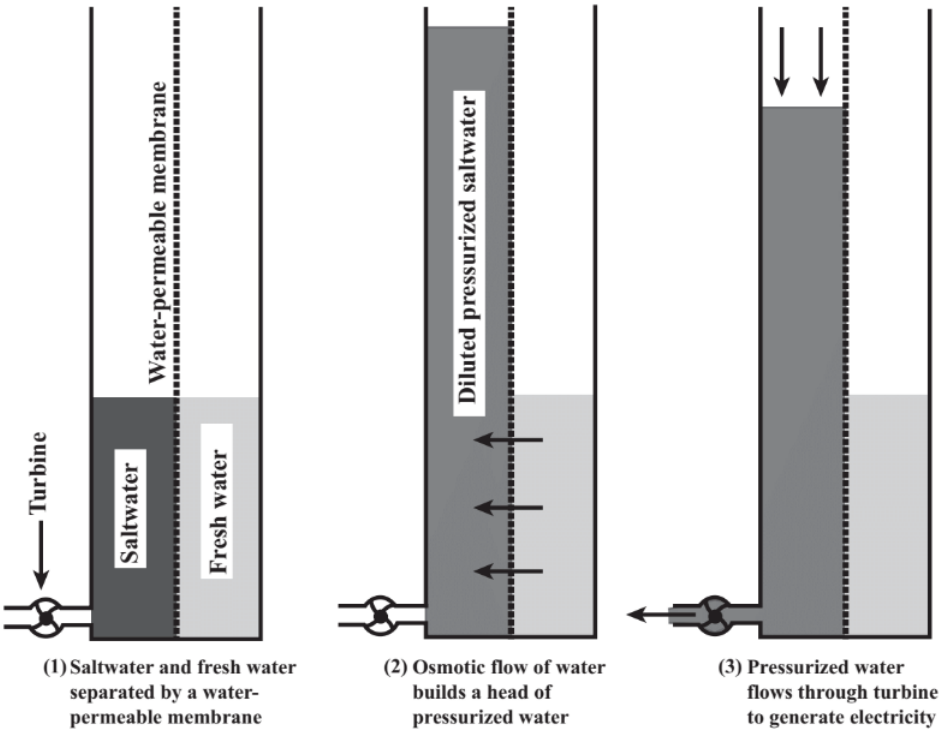


Figure 15.16. Illustration of the principle of pressure-retarded osmosis power generation that is based on the difference in osmotic pressure between saltwater and fresh water. Although it is shown here as a stepwise process, this mode of power generation operates on a continuous basis.

This page titled 2.5.15.12: Energy from Wind and Water is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.12: Energy from Wind and Water** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.13: Biomass Energy

Biomass and liquid and gaseous fuels made by processing it are the most promising sources of renewable energy for transportation.4As noted earlier in this chapter, photosynthesis,

$$6\,CO_2 + 6\,H_2O\ (\text{solar}\quad \text{energy}, h\nu)\ \rightarrow C_6H_{12}O_6\ (\text{glucose}\quad \text{carbohydrate}) + 6\,O_2$$

enables the conversion of solar energy to chemical energy in the form of biomass. Photosynthetically generated biomass supplies the food energy for essentially all organisms and,until about 200 years ago, was the source of most fuel. In addition to fuelwood and charcoal, **biofuels**, include forestry residues; agricultural byproducts; rapidly-growing grasses such as switchgrass; livestock manure; methane gas produced by anoxic fermentation of biomass; bioethanol from fermentation of sugar made from sugarcane, sugar beets, and cornstarch; and biodiesel synthesized from plant oils. As shown in Figure 15.6, photosynthesis suffers from the disadvantage of having less than 0.5% efficiency in the conversion of solar energy to chemical energy (although some plants, most notably sugarcane, convert solar energy to biomass energy with an efficiency of around 0.6%). Despite the limitations of photosynthesis in capturing energy, biomass is the predominant energy source in many developing regions of the world where approximately two billion people rely on wood for their primary household energy sources, especially for cooking. Today, more than 14% of the world's primary energy comes from biofuels, especially fuelwood and charcoal from wood. Finland gets a significant portion of its energy from burning the black liquor byproduct from pulp and paper manufacture, which uses wood as a raw material.

Properly utilized, biomass is a largely nonpolluting source of energy. Since it is produced by photosynthesis, there is no net addition to global atmospheric carbon dioxide. Although the heating value of dried biomass is only about half that of coal, biomass combustion produces very little sulfur dioxide, and the ash residue containing mineral nutrients can be returned to soil without adding harmful heavy metals, which can be a problem with coal ash.

The kinds of biomass that can be used for fuel fall into the four categories of (1) lignocellulosic materials from perennial plants, crop residues, wood, and biowastes; (2) starch from corn and other grains; (3) sugars from sugarcane and sugarbeets; and (4) oils from soybeans, rapeseed, and palm oil. These potential sources are addressed here.

As shown in Reaction 15.13.1, carbohydrates, such as glucose, $C_6H_{12}O_6$, are produced by photosynthesis. They can be burned directly, converted chemically to other fuels, or fermented to produce ethyl alcohol fuel. Hydrocarbons are more desirable as fuels, and some plants produce them directly. One example is the Philippine plant, *Pittsosporum reiniferum*, the fruits of which contain such a high content of hydrocarbon terpenes, primarily α-pinene and myrcene, that they can be burned to provide illumination. Rubber trees and other plants, such as *Euphorbia lathyrus* (gopher plant), a small bush growing wild in California, produce hydrocarbon emulsions. Seed oils, such as those produced by sunflowers and peanuts, and more exotic sources including buffalo gourd, cucurbits, and Chinese tallow tree, can be used for fuel, especially in diesel engines.

Despite concerns that not enough biomass can be grown to produce fuel and that it detracts from food supplies, it should be noted that about 150 billion metric tons of biomass are produced in the world each year by photosynthesis, mostly from uncontrolled plant growth. Corn, the most productive common field crop produces about 4 metric tons per acre of dry biomass annually(including stalks, leaves, husks, and corncobs). Switchgrass, a prolific producer of biomass, typically generates 11-12 tons of biomass per acre per year (there are 640 acres in a square mile of land). About 6% of the biomass generated globally each year would be equivalent to the world's demand for fossil fuels. Cultivation for fuel biomass of 6-8% of the land area of the 48 contiguous states would provide energy equivalent to annual U.S. consumption of petroleum and natural gas. Furthermore, only a small fraction of widely grown grain crops goes into grain; the rest is plant biomass, much of which could be used for energy production. And the U.S. has vast areas of underutilized land that could be devoted to the cultivation of energy-yielding plants. Much of this neglected, erosion-prone land would benefit from the cultivation of perennial plants that could be harvested for energy and regrow from roots left in the ground, thus lowering water and wind erosion.

A potentially important aspect of biomass fuel production that could improve its economics and environmental acceptability is coproduction of protein that can be used to feed animals. Now most protein used for feed comes from grain, particularly soybeans and corn. However, legumes(especially alfalfa) and perennial and annual grasses generate significant amounts of **leaf protein**. This material can be isolated by squeezing protein-rich juice from freshly ground plant leaves and heating the juice to precipitate the protein. The remaining fibrous material can then be used as a feedstock for the synthesis of biofuels.

Prolific Production of Biomass from Algae

Microscopic single-cell algae (microalgae) and photosynthetically-capable bacteria(cyanobacteria) growing in water can readily produce 10 times more biomass per unit area than terrestrial plants. In addition to their prolific productivity, microalgae offer several potential advantages for the generation of biofuels. These advantages include high production of oils and lipids (30-60% of dry algae mass), ability to grow in areas not suitable for terrestrial plants, ability to grow in saline water including seawater, growth in nutrient-rich sewage, and ability to grow in water enriched in dissolved carbon dioxide, such as from combustion sources. Concentrated by centrifugation and suspended in water or other liquid, microalgae can be introduced as a fluid emulsion into biorefineries offering processing advantages over soli.d biomass sources. Because of these advantages, microalgae are likely to become the predominant source of biofuels in the future.

Fuels from Fermentation of Biomass

Biological fermentation can be used to produce fuels from biomass. Yeasts act upon carbohydrates,

$$C_6H_{12}O_6 \rightarrow 2\,C_2H_5OH + 2\,CO_2$$

to produce ethanol, C_2H_5OH. This liquid alcohol can be used alone as a fuel, but is usually added to gasoline at levels of about 10% to produce **gasohol**, which burns more cleanly and with less CO output than ordinary gasoline. The source of carbohydrate for ethanol production is usually corn grain or sugar produced by sugarcane. Grain-based ethanol has seen strong growth in the U.S. and some other countries in recent years, much of it due to legislative initiatives. However, the net energy yield from this source is very low and its production competes with food crops, so it is nota very sustainable means of producing fuel. Ethanol from sugarcane sugars is significantly more competitive in Brazil, a prolific producer of sugarcane.

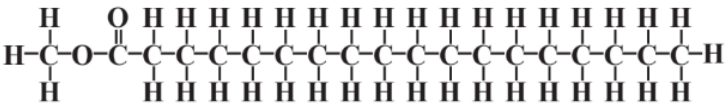
Another biomass fermentation occurs with methane-forming bacteria,

$$C_6H_{12}O_6 \rightarrow 3\,CH_4 + 3\,CO_2$$

to produce methane gas, CH_4. The gas mixture produced by this biochemical process can be burned directly, or the carbon dioxide can be removed to produce pure methane gas. Anoxic(oxygen-free) methane digesters used to degrade the biomass in sewage sludge, the residue from biological treatment of wastewater, can generate enough power to provide for the pumping and electrical needs of a large sewage treatment plant. Small methane digesters running on crop and food residues and human and animal wastes are used in rural areas of China to provide methane for cooking and lighting.

Biodiesel Fuel

Biodieselfuel is a combustible liquid synthesized from lipids, primarily those from plant oil seeds such as soybeans. Diesel fuel from petroleum consists of high-molecular-mass hydrocarbons containing 10-20 C atoms generally in a straight chain. Most lipids from plants are fatty acid esters of glycerol, such as triglyceride of stearic acid shown in Chapter 7, Figure 7.5. Examination of this formula shows that the fatty acid entities in it are predominantly straight-chain hydrocarbons except for the oxygen-containing carboxylate groups through which they are attached to the glycerol alcohol. Hydrolysis of the triglyceride and esterification with methanol yields liquid esters such as methyl stearate,



Methyl stearate, an ester in biodiesel fuel

that have the combustion characteristics needed for diesel fuel. In addition to methyl stearate, biodiesel fuel contains methyl esters of other fatty acids including linoleic acid, oleic acid, lauric acid, and behenic acid. Unlike ethanol, which cannot be transported through existing pipeline systems because it absorbs water, is therefore corrosive, and must be blended with gasoline at the point of distribution, biodiesel can be handled through existing facilities.

Biodiesel fuel can be synthesized from oils extracted from rapeseed (the major source in Europe), soybean (predominant in the U.S.), sunflower, palm, coconut, and jatropha. Both rapeseed and soybeans leave a protein-rich byproduct after the oil has been removed that is a good food source for animals. Palm oil, coconut oil and jatropha (from *Jatropha curcus*, planted for hedges) are

attractive oil sources for biodiesel production because they come from perennials that thrive in the tropics. Unfortunately, the explosive growth of palm oil tree plantations in Malaysia and Indonesia has resulted in high levels of rain forest destruction.

As noted above, microalgae are prolific producers of biomass and these organisms may have oil contents exceeding 50% making algae attractive for the production of biodiesel fuel. Whereas the annual production of biodiesel from soybeans may reach 200 liters per acre and from palm oil 2,500 liters per acre, optimistic projections for biodiesel production from algae are as much as 40,000 liters per acre annually. Furthermore, as noted above, algae can be produced on desert lands and in saltwater, thus not competing with food crops.

This page titled 2.5.15.13: Biomass Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.13: Biomass Energy** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.14: Geothermal Energy

The final form of renewable energy to be considered here is **geothermalenergy** in the form of steam, hot water, or hot rock that produce steam used in steam turbines. First harnessed for the generation of electricity at Larderello, Italy, in 1904, geothermal power has since been developed in Japan, Russia, New Zealand, the Phillipines, at the Geysers in northern California and especially in Iceland, an island nation that essentially rests upon a bed of steaming hot rock.

The best source of geothermal energy is subterranean dry steam which, unfortunately, is rather rare. Steam mixed with superheated water is much more common, with the byproduct water varying from very high purity to water laden with corrosive, scale-forming salts. Badly contaminated water is normally injected back into the hot rock formation from which it came to prevent contamination of surface water.

Hot, dry rocks can be used to produce steam from water injected into fractured rock formations. This source of geothermal energy is potentially ten times that of steam and hot water sources, but has been hindered largely over concerns raised by miniscule earthquakes that have resulted from fracturing the rocks. Development of this source continues on an experimental basis.

An interesting possibility that has yet to be demonstrated is the use of supercritical carbon dioxide as a working fluid for the extraction of energy from hot rocks (Figure 15.16).

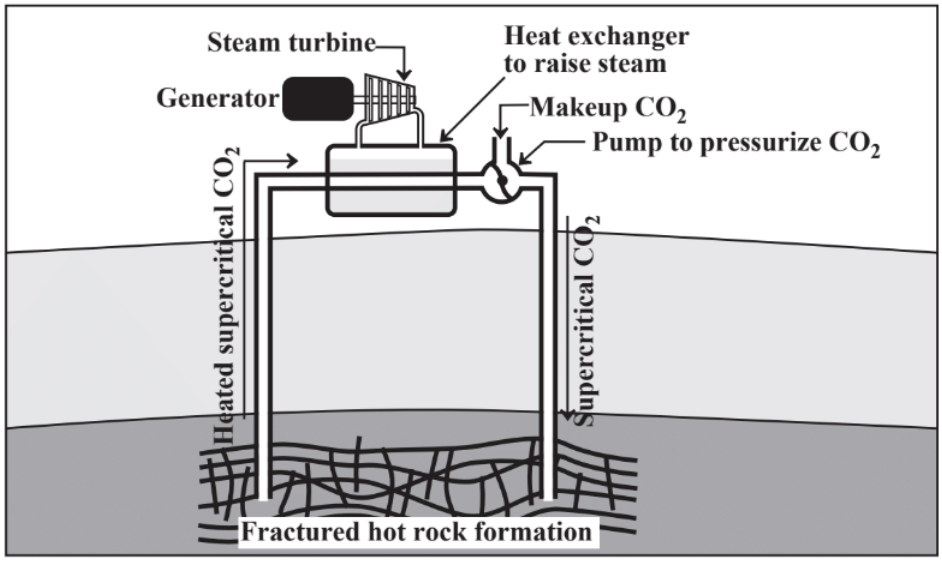


Figure 15.16. A system for extracting heat from hot dry rock formations using supercritical or dense phase carbon dioxide. The carbon dioxide is pumped into a fractured hot rock formation through an injection well and is removed through a production well. The hot carbon dioxide is run through a heat exchanger to raise steam to run a turbine that is coupled to an electrical generator and through which water and steam are recycled. Makeup carbon dioxide is added to the carbon dioxide line downstream from the heat exchanger and the carbon dioxide is repressurized and recycled through the hot rock formation.

This page titled 2.5.15.14: Geothermal Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.14: Geothermal Energy** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.15: Hydrogen for Energy Storage and Utilization

Hydrogen gas, H_2, can serve as a pollution-free means to store and utilize energy. Elemental hydrogen is the least polluting of fuels because it produces only water. Electricity that is generated on an intermittent basis from solar, wind, and even tidal flow processes can be used to electrolyze water,

$$2\,H_2O + \text{electrical}\quad \text{energy}\ \rightarrow 2\,H_2\,(g) + O_2\,(g)$$

and the hydrogen piped some distance and combusted in an engine, used in a fuel cell (Figure 15.5(5)), or stored by pumping it underground. Although not yet practical, direct photochemical processes that can split water molecules to H_2 and O_2 are attractive from the viewpoint of sustainability.

Hydrogen is employed to a limited extent to power vehicles. Such vehicles using internal combustion engines fueled with hydrogen are especially well adapted to Iceland where abundant geothermal and hydroelectric power provide ample supplies of electrolytically generated hydrogen. Furthermore, in Iceland it is not possible to drive long distances so that one is unlikely to be stranded far from a refueling station. Honda has made a very limited number of fuel-cell-powered automobiles available for lease in southern California where the company has established hydrogen fueling stations.

The idea of a “hydrogen economy” in which H_2 gas is the predominant medium of energy transfer, storage, and utilization may be too optimistic because of factors such as hydrogen’s low heating value per unit volume and the wide range of explosive mixtures it forms with air. One of the greatest barriers to the widespread adoption of hydrogen-fueled vehicles has been their inability to carry sufficient hydrogen for an acceptable range. Several solutions to this problem are now being investigated. One possibility is the storage of hydrogen in high-strength containers at up to almost 700 times atmospheric pressure reputed to contain sufficient hydrogen to propel an automobile 300 miles. Solids with very high sorptive capacities for H_2 (“super activated carbons”)are being investigated for hydrogen storage. Liquid fuels including gasoline and methanol can be broken down catalytically to generate hydrogen, but in so doing release greenhouse gas carbon dioxide.

Unlike fossil fuels such as methane, elemental hydrogen is not a primary source of energy and must be produced from other energy sources. In addition to generation by the electrolysis of water, most hydrogen now is produced by steam reforming of methane from natural gas

$$CH_4 + H_2O\ \rightarrow 3\,H_2 + CO$$

The carbon monoxide product can be reacted with steam,

$$CO + H_2O\ \rightarrow CO_2 + H_2$$

to produce additional H_2 and the CO_2. Such a process is counterproductive in providing fuel especially in that methane gas is easier to store and transport than elemental hydrogen and the modern internal combustion engine with associated emissions controls running on methane is virtually pollution-free. So the intermediate step to produce elemental hydrogen is not a very sustainable approach. Although production of elemental hydrogen by electrolysis of water using electricity from renewable sources as discussed above is essentially non-polluting, electrolysis is a relatively inefficient means of using electrical energy, which might more sustainably be used for example in charging batteries in plugin hybrid vehicles.

This page titled 2.5.15.15: Hydrogen for Energy Storage and Utilization is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **15.15: Hydrogen for Energy Storage and Utilization** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.15.16: Literature Cited and Supplementary References

LITERATURE CITED

1. “Fuel Cell Technologies Program,” U.S. Department of Energy, 2010,www1.eere.energy.gov/hydrogen.../fc_types.html.

2.Spicher, U., and T. Heidenreich, “Stratified-Charge Combustion in Direct Injection Gasoline Engines,” *Advanced Direct Injection Combustion Engine Technologies and Development*, Vol.1, 2010, pp. 20-44.

3. Balban, Richard C., Josephine A. Elia, and Christodoulos A. Floudas, “Toward Novel Hybrid Biomass, Coal, and Natural Gas Processes for Satisfying Current Transportation Fuel Demands: Process Alternatives, Gasification Modeling, Process Simulation, and Economic Analysis,” *Industrial and Engineering Chemistry Research*, 7343-7370 (2010).

4. “The Role of Biomass in America's Energy Future,” *Biofuels, Bioproducts and Biorefining* (special issue), **3**, 105-288 (2009)

SUPPLEMENTARY REFERENCES

Al-Hallaj, Said, and Kristofer Kiszynski, *Hybrid Hydrogen Systems: Stationary and Transportation Applications*, Springer, New York, 2010.

Ball, Michael, and Martin Wietschel, Eds., *The Hydrogen Economy: Opportunities and Challenges*, Cambridge University Press, Cambridge, UK, 2010.

Boxwell, Michael, *Solar Electricity Handbook, 2010 Edition: A Simple Practical Guide to Solar Energy - Designing and Installing Photovoltaic Solar Electric Systems*, 3rd ed., Greenstream Publishing, Warwickshire, UK, 2010.

Brebbia, C. A., and V. Popov, Eds., *Energy and Sustainability*, WIT Press, Southampton, U.K.,2007.

Coley, David A., *Energy and Climate Change: Creating a Sustainable Future*, Wiley, Hoboken,NJ, 2008.

Davis, Scott, *Serious Microhydro: Water Power Solutions from the Experts*, Gabriola Island, British Columbia, New Society Publishers, 2010.

Hefner, Robert A., *The Grand Energy Transition: The Rise of Energy Gases, Sustainable Life and Growth, and the Next Great Economic Expansion*, Wiley, Hoboken, NJ 2009.

Herbst, Alan M., and George W. Hopley, *Nuclear Energy Now: Why the Time Has Come for the World’s Most Misunderstood Energy Source*, Wiley, Hoboken, NJ, 2007.

Hofman, Konrad A., Ed., *Energy Efficiency, Recovery and Storage*, Nova Science Publishers, New York, 2007.

Huenges, Ernst, and Patrick Ledru, Eds., *Geothermal Energy Systems: Exploration, Development and Utilization*, Wiley-VCH, Hoboken, NJ, 2010.

Infield, David, and Leon Freris, *Renewable Energy in Power Systems*, John Wiley & Sons, Chichester, U.K., 2008.

Krauter, Stefan C. W., *Solar Electric Power Generation — Photovoltaic Energy Systems*, Springer, New York, 2006.

Mahaffey, James, *Atomic Awakening: A New Look at the History and Future of Nuclear Power*, Pegasus, Cambridge, UK, 2010.

Manwell, James.F., Jon G. McGowan, and Anthony L. Rogers,*Wind Energy Explained: Theory, Design and Application*, 2nd ed., Wiley, Hoboken, NJ, New York, 2010.

Nag, Ahinda, Ed., *Biofuels Refining and Performance*, McGraw-Hill, New York, 2008.

Nakaya, Andreas, *Energy Alternatives*, Reference Point Press, San Diego, CA, 2008.

Romm, Joseph J., *The Hype About Hydrogen: Fact and Fiction in the Race to Save the Climate*, Island Press, Washington, 2004.

Soetaert, Wim, and Erick Vandamme, *Biofuels*, Wiley, Hoboken, NJ, 2008.

Suppes, Galen J., and Truman S. Storvick, *Sustainable Nuclear Power*, Elsevier/Academic Press, Amsterdam, 2007.

Tabak, John, *Solar and Geothermal Energy*, Facts On File, New York, 2009.

Vertes, Alain, Nasib Qureshi, Hideaki Yukawa, and Hans Blashek, Eds.,*Biomass to Biofuels: Strategies for Global Industries*, Wiley, Hoboken, NJ, 2010.

Weiss, Charles, and William B. Bonvillian, *Structuring an Energy Technology Revolution*, MIT Press, Cambridge, MA, 2009.

Worldwatch Institute, *Biofuels for Transport: Global Potential and Implications for Energy and Agriculture*, Earthscan, Sterling, VA, 2007.

2.5.15.16: Literature Cited and Supplementary References is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Literature Cited and Supplementary References** has no license indicated.

2.5.15.17: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. The tail of a firefly glows, although it is not hot. Explain the kind of energy transformation that is most likely involved in the firefly's producing light.

2. What is the standard unit of energy? What unit did it replace? What is the relationship between these two units?

3. Which law states that energy is neither created nor destroyed.

4. What is the special significance of 1,340 watts?

5. What is the reaction in nature by which solar energy is converted to chemical energy?

6. In what respects is wind both one of the oldest, as well as one of the newest, sources of energy?

7. What are two major problems with reliance upon coal and petroleum for energy?

8. Why does natural gas contribute less to greenhouse warming than does petroleum and much less than coal?

9. How might coal be utilized for energy without producing greenhouse gas carbon

10. What is a large limiting factor in growing biomass for fuel, and in what respect does this limit hold hope for the eventual use of biomass fuel?

11. What relationship describes the limit to which heat energy can be converted to mechanical energy?

12. Why does a diesel-powered vehicle have significantly better fuel economy than a gasoline-powered vehicle of similar size?

13. Why is a nuclear power plant less efficient in converting heat energy to electricity than is a fossil-fueled power plant?

14. Instead of having a sparkplug that ignites the fuel, a diesel engine has a glow plug that operates only during engine startup. Explain the operation of the glow plug

15. Cite two examples of vastly increased efficiency of energy utilization that took place during the 1900s.

16. Describe a combined power cycle. How may it be tied with district heating?

17. Describe a direct and an indirect way to produce electricity from solar energy.

18. What is the distinction between donor and acceptor layers in photovoltaic cells?

19. Using internet resources for information list some possible means for storing energy generated from solar radiation?

20. What are the advantages of *Pittsosporumreiniferum*and*Euphorbialathyrus* for the production of biomass energy?

21. Corn produces biomass in large quantities during its growing season. What are two potential sources of biomass fuel from corn, one that depends upon the corn grain and the other that does not?

22. Does the use of biomass for fuel contribute to greenhouse gas carbon dioxide? Explain.

23. What fermentation process is used to generate a fuel from wastes, such as animal wastes?

24. What are two potential pollution problems that accompany the use of geothermal energy to generate electricity?

25. What basic phenomenon is responsible for nuclear energy? What keeps the process going?

26. What is the biggest problem with nuclear energy? Why is it not such a bad idea to store spent nuclear fuel at a reactor site for a number of years before moving it?

27. What is meant by passive stability in nuclear reactor design?

28. What is the status of thermonuclear fusion for power production?

29. Arrange the following energy conversion processes in order from the least to the most efficient:(a) electric hot water heater, (b) photosynthesis, (c) solar cell, (d) electric generator, (e) aircraft jet engine.

30. Considering the Carnot equation and common means for energy conversion, what might be the role of improved materials (metal alloys, ceramics) in increasing energy conversion efficiency?

31. Justify describing the sun as “an ideal energy source.” What are two big disadvantages of solar energy?

32. What are some of the greater implications of the use of biomass for energy? How might such widespread use affect greenhouse warming? How might it affect agricultural production of food?

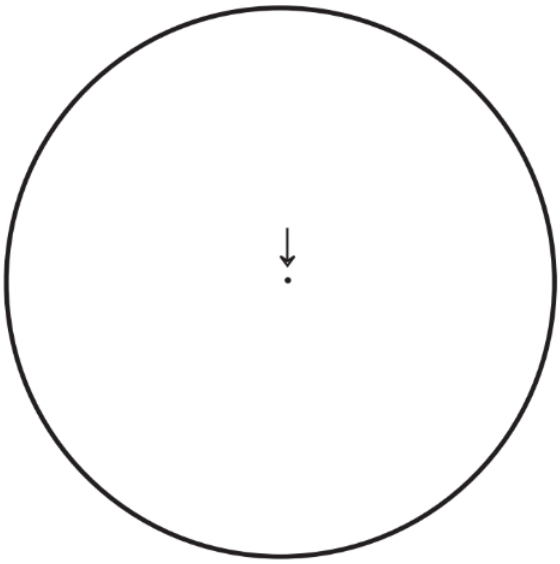
2.5.15.17: Questions and Problems is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** has no license indicated.

SECTION OVERVIEW

2.5.16: Terrorism, Toxicity, and Vulnerability- Green Chemistry and Technology in Defense of Human Welfare

“If the area of the large circle below represents a fatal dose of the once-popular organophosphate insecticide parathion, now banned because of its toxicity, the area of the barely perceptible dot indicated by the arrow represents a fatal dose of organophosphate Sarin a nerve gas military prison of concern for its potential use in terrorist attacks.”



2.5.16.1: Vulnerability to Terrorist Attack

2.5.16.2: Protecting the Anthrosphere

2.5.16.3: Substances That Explode, Burn, or React Violently

2.5.16.4: Toxic Substances and Toxicology

2.5.16.5: Toxic Chemical Attack

2.5.16.6: Protecting Water, Food, and Air

2.5.16.7: Detecting Hazards

2.5.16.8: Green Chemistry to Combat Terrorism

2.5.16.9: Green Chemistry for Sustainable Prosperity and a Safer World

2.5.16.10: Green Chemistry to Combat Terrorism

2.5.16.11: Green Chemistry for Sustainable Prosperity and and a Safer World

2.5.16.12: Literature Cited

2.5.16.13: Questions and Problems

This page titled 2.5.16: Terrorism, Toxicity, and Vulnerability- Green Chemistry and Technology in Defense of Human Welfare is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.16.1: Vulnerability to Terrorist Attack

Terrorist attacks upon human targets have become a constant fear in modern times. In the United States, vulnerability to such attacks were illustrated in horrifying detail by the suicide attacks by hijacked commercial aircraft on the New York World Trade Center on September 11, 2001. Other nations have long lived in the shadow of threats from groups that would cause them harm. Throughout the world, the possibility of deliberate attacks upon people, their support systems, and the anthrospheric infrastructure have come to be the greatest concern facing large numbers of people.

Chemicals and chemistry figure prominently in considerations of terrorist actions. The sudden release of a huge amount of chemical energy from a mixture of ammonium nitrate (a common agricultural fertilizer) and diesel fuel brought down the Alfred P. Murrah Oklahoma City Federal Building in 1995 with the loss of dozens of lives. Powerful explosives strapped to the bodies of suicide bombers have killed 20 or more people at a time in attacks in Afghanistan, Iraq, Pakistan, and Israel. The extreme toxicity of military poison nerve gases is a constant concern in subways and other locations where large numbers of people are packed into small spaces. Biochemistryapplied to recombinant DNA science may enable production of particularly virulent disease pathogens, such as vaccine-resistant smallpox. The accidental release of methyl isocyanate in an industrial chemical accident in Bhopal, India, in 1984 killed more people than even the 2001 attack on the World Trade Center. At least 243 people died from hydrogen sulfide contained in natural gas released from a pressurized deposit of this lethal mixture penetrated by a drilling operation in the Chuandongbei natural gas field of southwestern China in December, 2003. Hundreds of people were made ill and thousands were evacuated. A massive fire resulted when the escaping gas was ignited to convert the hydrogen sulfide (H_2S) to toxic, but much less lethal sulfur dioxide, SO_2.

Terrorist activities are not confined to direct attacks upon humans. The environment is susceptible to terrorist activities and may be severely damaged by them. For example, a major nuclear war — arguably the ultimate form of terrorism — could contaminate large areas of land and other parts of the environment with radioactive materials and, in the worst case scenario, could do substantial harm to the global climate resulting in a “nuclear winter.”

So, what can green chemistry do to prevent terrorist attacks and mitigate their effects? Actually, green chemistry is a key discipline in such endeavors. For example, one of the basic tenets of green chemistry is to use the safest possible chemicals as safely as possible. When particularly dangerous chemicals are not made or used, they are not available to cause mischief. The practice of green chemical manufacturing calls for minimizing the accumulation of hazardous chemicals and seeks to eliminate hazardous chemical wastes. Safer materials made under the practice of green chemical technology minimize hazards from more dangerous substances. Highly sensitive analytical techniques developed by chemical science can be used to detect minuscule quantities of explosives or toxic substances slated for use in terrorist attacks. Biochemistry and recombinant DNA science have the potential to enable the development of better vaccines against pathogenic biological warfare agents and antidotes to chemical and biological toxins. More subtly, the use of green chemistry and chemical technology to produce effective substitute materials can reduce potential for “resource blackmail” that can lead to vulnerability to terrorist activity. A prime example is the substitution of biomass alternatives for petroleum feedstocks that to a certain extent many nations must obtain from potentially unfriendly nations.

This chapter addresses potential terrorist threats with emphasis upon those that employ chemical and biological agents. Having identified threats that may occur, it then discusses ways in which chemistry, especially the proper practice of green chemistry, can minimize such threats.

This page titled 2.5.16.1: Vulnerability to Terrorist Attack is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **16.1: Vulnerability to Terrorist Attack** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.16.2: Protecting the Anthrosphere

The anthrosphere constructed with a high degree of human ingenuity has provided a generally safe and comfortable environment for human beings. The underpinning of this entire support system is the **infrastructure**, which includes systems to purify and deliver water, electricity generation and distribution systems, communications, fuel distribution networks, highways, and railroads. The sophistication and complexity of the infrastructure is what makes it work so smoothly (for the most part), but also makes it vulnerable to attack. A key aspect of this susceptibility to attack is **vulnerability due to interconnectivity**, which arises from the high degree to which various parts of the infrastructure are interconnected and mutually dependent. No part of the infrastructure illustrates vulnerability due to interconnectivity more so than modern electrical grids, which can stretch across vast geographical regions and across national borders. A reminder of this vulnerability occurred on August 14, 2003, when a failure of the electrical grid in the northeastern U.S. and southeastern Canada resulted in a power outage for tens of millions of people in New York City, Detroit, Cleveland, and Toronto. In this incident a total of 68,100 megawatts of generating capacity — equivalent to 68 very large, modern power generating facilities — was lost as dozens of high-voltage transmission lines shut down. This occurred within about 5 minutes, and the event that caused it probably took only about 10 seconds. The precipitating event was not terrorism — although it well could have been — and probably resulted from nothing more dramatic than tree limbs interfering with transmission lines.

The electrical power failure described above illustrates a phenomenon called **cascading failures** on complex networks discussed in a 2009 article in *Scientific American*.1An electrical power distribution system is a complex network of hundreds of electrical power plants interconnected by electrical transmission lines. Such a system enables very efficient generation, distribution, and utilization of electrical power in that a surge in demand on one part of the system can be compensated by redistribution of power from the network to that segment of the system. Therefore, local generating facilities do not have to maintain the capacity to meet temporary demand, and this results in high efficiency and much lower costs. Other parts of the infrastructure have similar networks. Routers on the Internet are designed to route Internet traffic around bottlenecks or other routers that are temporarily out of commission. Modern manufacturing operations using “just-in-time” deliveries of components make it unnecessary to stockpile large quantities of parts, thereby minimizing the need for production capacity and maximizing efficiency. The downside is that these systems operate “close to the edge” so that a relatively small failure, such as one brought about by human mischief, can rapidly cascade into a major failure.

Arguably the part of the infrastructure most vulnerable due to interconnectivity and subject to cascading failures resulting from terrorist attack is the vast, intricately interconnected computer network that is now part of all modern systems of communication, commerce, and military operations.[2] These systems are vulnerable to so-called cyber crime in which key components can be disrupted leading in worst cases to total system breakdown. Much of the effort in homeland security in the U.S. and similar efforts in other countries is devoted to combatting cyber crime.

Chemistry can be applied to infrastructure protection. One area in which this is true is the production of materials that resist heat and flame. Such materials used in buildings can provide substantial protection from fire. Processes that are consistent with the practice of green chemistry also provide protection from attack. For example, green chemistry attempts to reduce the production and use of hazardous materials. Sophisticated analytical chemistry and analytical instrumentation can be used to detect agents of attack before damage is done.

This page titled 2.5.16.2: Protecting the Anthrosphere is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **16.2: Protecting the Anthrosphere** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.16.3: Substances That Explode, Burn, or React Violently

Substances that react violently and rapidly enough to cause fires, explosions, or corrosive destruction of materials are those that have been used most commonly in terrorist attacks and that have the most potential for future use. Reactive substances, such as explosives used to quarry rock, have many commercial uses and are therefore widespread and potentially readily available, both to legitimate users and terrorists.

Explosives are the most common materials used in terrorist attacks. The devastating effects of explosives have been illustrated by a number of incidents including the 1995 bombing of the Murrah Federal Building in Oklahoma City, the 2003 bombing of the British consulate in Turkey,and frequent bombings of various sites in Iraq in 2003/2004. Explosives can be made from readily available materials; the Murrah Federal Building was brought down by a mixture of ammonium nitrate fertilizer and diesel fuel. Gunpowder has long been prepared from charcoal, sulfur, and nitrate salts. Nitroglycerin, the explosive in dynamite, is made from glycerin, a byproduct of soap preparation, reacted with nitric and sulfuric acids (most amateurs who attempt this synthesis succeed in blowing their heads off). More sophisticated explosives consist of organic compounds containing nitrogen and oxygen, such as 2,4,6-trinitrotoluene (TNT), 1,3,5-trinitro-1,3,5-triazacyclohexane (RDX), and pentaerythritol tetranitrate (PETN). Figure 16.1 gives structural formulas of several explosives that could be used by terrorists.

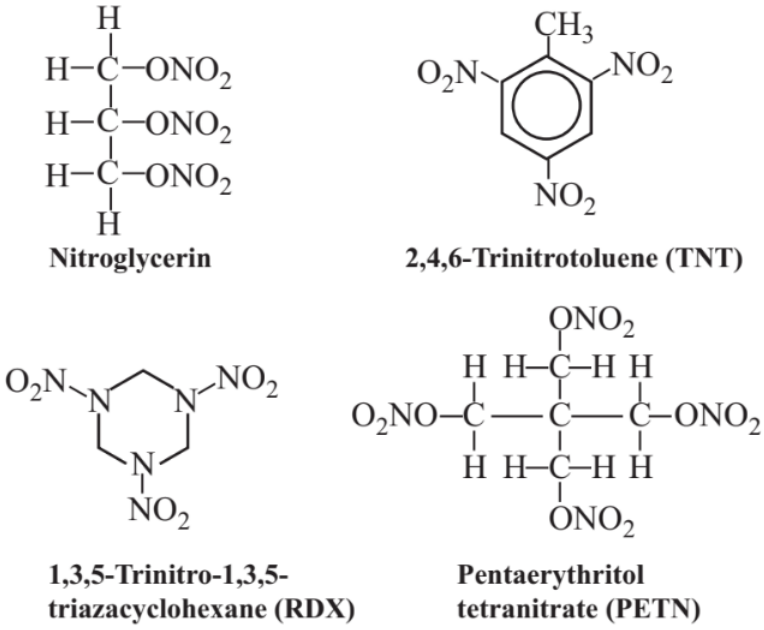


Figure 16.1. Structural formulas of four explosive compounds that have the potential for use in terrorist attacks. Note that all contain an abundance of oxygen bound with nitrogen.

The potential of flammable substances to cause death and destruction was shown by the jet-fuel-fed fires that brought down both towers of the New York World Trade Center in the 9/11/01 attack. Fires set on ruptured petroleum pipelines have caused significant destruction in Iraq in 2003/2004. When mixed with air, vapors of flammable liquids can cause massive explosions. Adding to the hazard of flammable substances are oxidants that release oxygen to react with combustible materials. In 1997, oxidant sodium chlorate, $NaClO_3$, an ingredient of emergency oxygen generators in aircraft that were being transported contrary to regulations in the cargo hold of a Valujet airplane, caused a devastating fire of aircraft tires that brought the aircraft down in the Florida Everglades.

By their nature, fuels are flammable substances, as are some common industrial solvents. Flammable fuels and solvents are used throughout almost all societies and are readily available to those who would use them for malevolent ends. Adding to their availability is their transport by truck, rail, and pipeline. The hazard of flammable substances is increased by the ease by which they and their vapors can be distributed through sewers, elevator shafts, subway tunnels, and other conduits.

Corrosive substances that destroy materials and flesh have been used in attacks on equipment and people. Concentrated sulfuric acid, which dehydrates and destroys flesh, has been used by criminals to blind people. Corrosive materials can be used to damage relays and other devices in communications equipment as a means of sabotage.

A major concern with respect to flammable, reactive, and explosive substances is their widespread industrial use. Actually, such materials are relatively safe inside of manufacturing plants and properly secured storage areas. The greater threat comes from their transport. This is illustrated by very frequent transportation accidents involving rail cars, trucks, barges, and pipelines that result in explosions, fires, and release of corrosive materials. Hijacking of trucks transporting hazardous materials and even trucks driven by terrorists are a particular concern. The practice of industrial ecology and green chemistry can help minimize such threats by, for example, promoting the production of hazardous substances in minimal quantities where needed and as needed. “Just-in-time” production minimizes storage of hazardous substances.

This page titled 2.5.16.3: Substances That Explode, Burn, or React Violently is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **16.3: Substances That Explode, Burn, or React Violently** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.16.4: Toxic Substances and Toxicology

One of the greater concerns that the general public has with chemistry is the potential toxic effects of various substances including those that could be used for terrorist attacks. **Poisons**, or **toxicants**, are substances that can adversely affect biological tissue leading to harmful responses including, in the severest cases, even death. The study of such substances and their effects is the science of **toxicology**. The science that relates the chemical properties of toxic substances to their toxic effects is **toxicological chemistry**. Because poisons are among the leading terrorist threats, it is appropriate to consider toxic substances and toxicological chemistry here.

Any kind of tissue and all organs can be the subject of attack by toxic substances. The major human organ systems that are potentially adversely affected by toxic substances are given in Table 2.5.16.4.1.

Table 2.5.16.4.1

Table 2.5.16.4.1	Typical Toxic Responses
Respiratory system	Emphysema from cigarette smoke, lung cancer from asbestos
Skin responses	Allergic contact dermatitis, such as from exposure to dichromate; chloracne from exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin ("dioxin"); skin cancer from exposure to coal tar constituent
Hepatotoxicity (toxic effects)	Steatosis (fatty liver), such as from exposure to carbon tetrachloride cirrhosis; (deposition and build up of fibrous collagen tissue) from excessive ingestion of ethanol; haemangiosarcoma, a type of liver cancer observed in workers heavily exposed to vinyl chloride in PVC plastic manufacture
Reproductive system	Interference with sperm development by some industrial chemicals, interference with cells involved with egg formation by chemicals such as cyclophosphamide
Blood	Carboxyhemoglobin formation from binding of carbon monoxide to blood hemoglobin, methemoglobinemia consisting of conversion of iron(II) to iron(III) in hemoglobin from exposure to substances such as aniline or nitrobenzene, aplastic anemia from exposure to benzene
Immune system effects	Immunosuppression from exposure to radiation, hypersensitivity from exposure to beryllium
Endocrine system effects	Disruption of endocrine function by endocrine disruptors such as bisphenol-A
Nervous system	Encephelopathy (brain disorder), such as from exposure to lead; peripheral neuropathy from exposure to organic solvents; inhibition of acetylcholinesterase enzyme in nerve function by exposure to organophosphate military poisons
Kidney and urinary tract system	Nephrotoxicity to the kidney by heavy metal cadmium

Toxicities

The toxicities of substances vary over a wide range, and those that are toxic at lowest doses are of most concern for deliberate poisoning. This is illustrated in Figure 16.1 which gives the toxicities of several substances. It is important to note that the dosage scale in this figure is logarithmic; that is for each division *decrease* on the scale, a substance is *tentimes* as toxic. The two circles in Table 2.5.16.4.1 illustrate the enormous differences between toxicities of different substances. If the area of the large circle represents the size of a fatal dose of parathion, a once widely used insecticide that has killed a number of people and has now been

banned because of its toxicity, a fatal dose of military poison nerve gas Sarin is represented by the minuscule dot below the circle! Toxicities are normally expressed as LD50 values, the dose in units of mass of poison per unit mass of test subject. Rats are usually used for tests, and toxicities to humans are inferred from these test values.

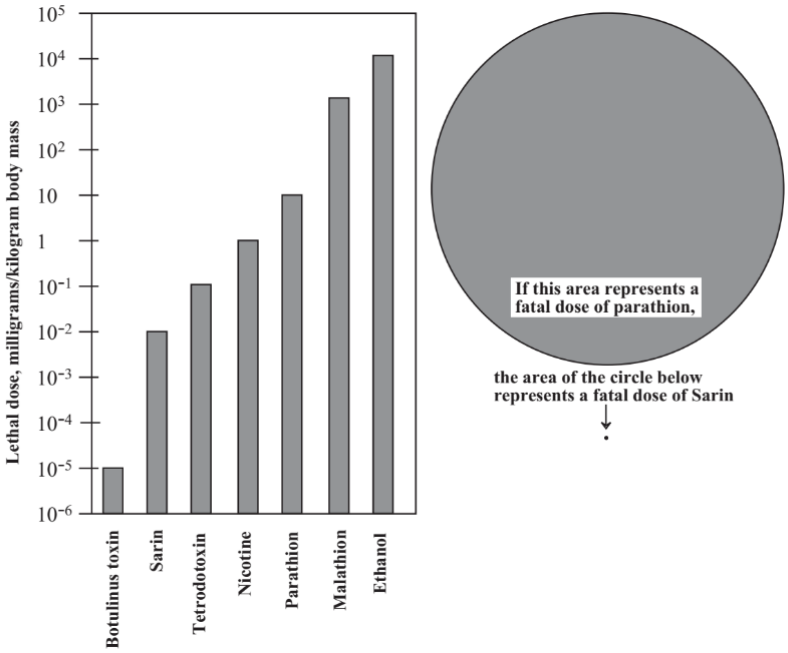
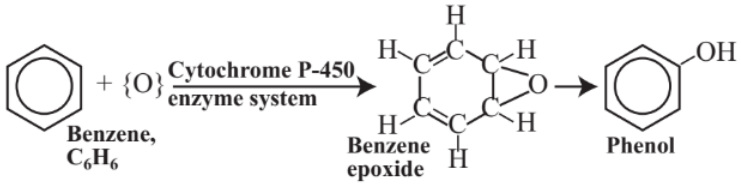


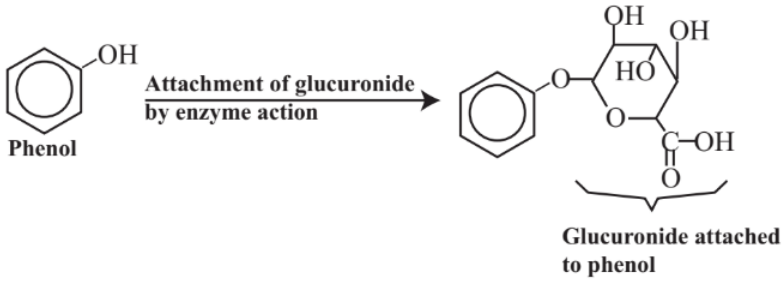
Table 2.5.16.4.1: Relative toxicities of various substances in units of milligrams dose per kilogram of body mass required to kill 50% of test subjects (LD50). Note that the estimated lethal dose is plotted on a log scale. Also shown are the relative toxicities of parathion, a once commonly used insecticide that has been discontinued because of its toxicity to humans, and Sarin, an extremely toxic military poison.

Metabolism of Toxic Substances

Toxic substances that enter the body and that are foreign to it, commonly called **xenobiotic substances**, are subject to metabolic processes that may activate them or make them less toxic(detoxification). The metabolism of toxic substances may be divided into two phases. Phase I reactions normally consist of attachment of a functional group, usually accompanied by oxidation. For example, benzene, C_6H_6, (see Chapter 6, Section 6.2) is oxidized in the body by the action of the cytochrome P-450 enzyme system as shown in Reaction 16.4.1. The Phase I oxidation product is phenol, a toxic substance. A reactive intermediate in the process is benzene epoxide, which interacts with biomolecules to cause toxic effects. The phenol Phase I oxidation product of benzene may undergo a second reaction, a Phase II reaction in which it is bound with a conjugating agent that is endogenous to (produced naturally by) the body, such as glucuronide as illustrated in Reaction 16.4.2



(16.4.1)



(16.4.2)

Although Phase I and Phase II reactions generally act to make xenobiotic substances more water soluble, more readily eliminated from the body, and less toxic, in some cases, the opposite occurs and metabolic processes make substances more toxic. Most known human carcinogens (cancer-causing agents) are actually produced by biochemical processes in the body from noncarcinogenic precursor substances.

The Action of Toxic Substances

Toxic substances, which, as noted above, are often produced by metabolic processes from nontoxic precursors, produce a toxic response by acting upon a **receptor** in the body. Typically, a receptor is an enzyme that is essential for some function in the body. As a consequence of the binding of the receptor to the toxicant there is a **biochemical effect**. A common example of a biochemical effect occurs when a toxicant binds to an enzyme such that the bound enzyme may be inhibited from carrying out its normal function. As a result of a biochemical effect, there is a response, such as a behavioral or physiological response, which constitutes the actual observed toxic effect. Acetylcholinesterase enzyme inhibited by binding to nerve gas Sarin may fail to stop nerve impulses in breathing processes, leading to asphyxiation. The phenomena just described occur in the **dynamicphase** of toxicant action as summarized in Figure 16.3.

This page titled 2.5.16.4: Toxic Substances and Toxicology is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **16.4: Toxic Substances and Toxicology** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.16.5: Toxic Chemical Attack

Other than nuclear weapons, chemical and biological agents (pathogenic microorganisms) have more potential than any other mode of attack to debilitate and kill victims. Not the least of these effects is their ability to spread fear and terror among potential victims. Most of what is known of the potential of chemical and biological agents to harm and terrorize people has been learned from their use and preparation for use in warfare as summarized in a book on the history of weapons of terror.[3]

Chemical and biological agents have been used for centuries by civilizations around the world that have poisoned water supplies and employed diseased animals and human bodies, incendiary materials, poison-tipped arrows, and even venomous snakes to attack their enemies. Modern chemical warfare normally is considered to date from the use of toxic chlorine gas by the German army in an attack at Ypres, Belgium, during World War I in 1915 causing 7,000 to 15,000 casualties. The British used chlorine five months later in the Battle of Loos, but suffered 2,000 casualties when a change of wind blew the gas back over their own lines. The use of chemical agents continued throughout the war and, in addition to physical damage that was done, contributed to psychological stress and the need to implement cumbersome protective measures and logistics. The Geneva Protocol of 1925 banned the use of chemical and biological agents in warfare, although before, during, and after World War II nations continued development of these means of attack.

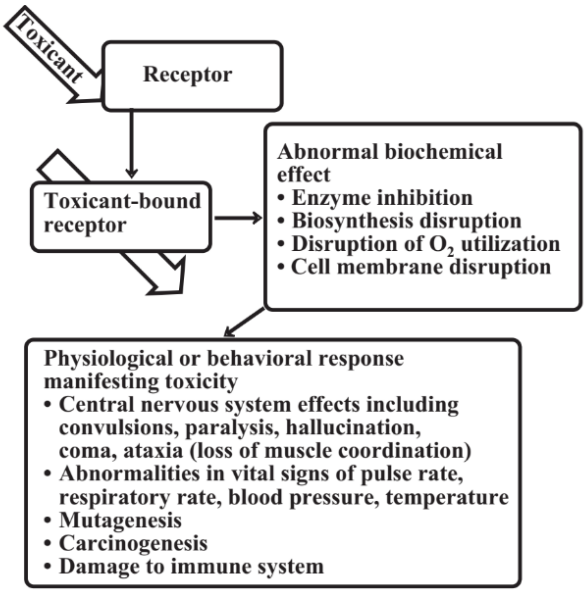


Figure 2.5.16.5.1: The dynamic phase of toxicity in which a toxicant interacts with a receptor in the body to cause a biochemical alteration that is manifested in a toxic response

The next large scale use of chemical weapons took place in the 1980-1988 war between Iraq and Iran in which Iraq employed approximately 1,800 tons of blistering agent mustard gas and 140 and 600 tons respectively of nerve gas sarin and Tabun causing an estimated 30,000 casualties. Toward the end of this conflict Iraq killed and injured thousands of people with chemical agents in quelling a Kurdish rebellion. Fear of missile-launched chemical attack by Iraq was a major concern during the war that followed Iraq's 1990 invasion of Kuwait. Suspected possession by Iraq of such "weapons of mass destruction" was cited as justification for the 2003 invasion of Iraq by the U.S.

Biological warfare agents are regarded as having significantly more potential than chemical agents for inflicting casualties. This is because of the ability of pathogens to propagate and afflict far more people than those exposed in the initial attack.

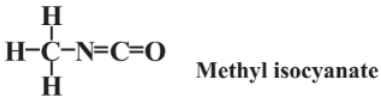
Chemical and biological agents are probably more effective for terrorist attacks against civilians than they are for warfare. Civilian targets are generally less likely to have protective measures in place, are more subject to surprise attack, and can be attacked in enclosed areas with relatively unsophisticated means. As exemplified by the October 2001 anthrax bacteria attack on several targets carried out through the U.S. Postal Service, which caused 22 cases of the disease and killed 5 people, panic fueled in part by intense media attention can cause widespread panic and disruption, always the goal of a terrorist attack.

This page titled 2.5.16.5: Toxic Chemical Attack is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **16.5: Toxic Chemical Attack** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.16.6: Protecting Water, Food, and Air

The tragedy that arguably illustrates most vividly the potential of chemicals to cause major loss of life was the accidental release of methyl isocyanate from a chemical manufacturing operation in Bhopal, India. This disaster occurred during the night of December 2/3, 1984, exposing many victims as they slept in their homes. A total of about 40 tons of this chemical was released during the incident exposing thousands of residents in the surrounding area. Of those exposed, more than 3000 died, primarily from pulmonary edema (fluid accumulation in the lung). Immunological, neurological, ophthalmic (eye), and hematological effects were also observed. Because of its high vapor pressure and toxicity to multiple organs,

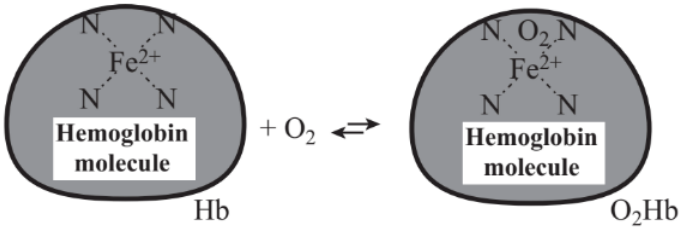


methyl isocyanate is the most toxic of the isocyanates. An interesting aspect of methyl isocyanate toxicity is its ability to cross cell membranes and reach organs far from the site of exposure, despite its very high reactivity.

It is unlikely that methyl isocyanate or of any other substance with similar toxicity could be obtained and delivered to potential victims in large enough quantities and with sufficiently effective delivery to cause widespread death and illness in a deliberate attack. Nevertheless, the magnitude of the Bhopal catastrophe vividly illustrates the potential of toxic substances to cause harm and the incident serves as a grim reminder of the potential of toxic chemicals, especially those that may be airborne, to be used in attacks. And the Bhopal incident, which occurred as the result of a tragic accident, serves as a grim warning of the potential of terrorist attacks upon chemical plants to cause widespread harm.

Potential Chemical Agents

Although an unlikely "weapon of mass destruction," **carbon monoxide**, **CO**, can certainly cause fatal poisonings. It has killed thousands of people accidentally and through suicide. Given the fact that carbon monoxide is odorless and provides essentially no warning of exposure, it should be regarded as a potential weapon. To understand the action of carbon monoxide, it is important to realize that oxygen in blood is carried from the lungs to tissue by **hemoglobin**, a large-molecule protein in red blood cells that contains iron(II) (the Fe^{2+} ion) bound with nitrogen. This functionality exchanges oxygen



converting hemoglobin, Hb, into oxyhemoglobin, O_2Hb, which carries oxygen to tissues where it is released to be used for metabolic processes. When carbon monoxide is present in inhaled air, the oxygen bound to hemoglobin is displaced,

$$O_2Hb + CO \rightleftharpoons COHb + O_2$$

producing **carboxyhemoglobin**, COHb. This species is not only useless for carrying oxygen, it is much more stable than oxyhemoglobin, O_2Hb, so that a relatively low concentration of carbon monoxide will convert enough of the hemoglobin to carboxyhemoglobin to cause a serious oxygen deficiency. Rapid death ensues from inhalation of air containing 1,000 parts per million (ppm) carbon monoxide, and unconsciousness results from inhaling 250 ppm CO. Dizziness, headache, and fatigue result from inhalation of 100 ppm CO and levels as low as 10 ppm can impair visual perception and judgment.

Chlorine (Cl_2) gas could potentially be used in terrorist attacks because of its wide availability for water disinfection and other uses. Illustrative of this potential is the fact that chlorine was the first substance used as a military "poison gas" in World War I. Chlorine is a strong oxidizer that reacts with water, including water in tissue, to produce an oxidizing, acidic solution that is especially damaging to respiratory (lung) tissue. Air containing only 10–20 ppm chlorine causes acute discomfort to the respiratory

tract and brief exposure to 1,000 ppm of Cl_2 can be fatal. Because of its widespread use and transport as the liquid in railway tank cars, chlorine is regarded as having a high potential as a terrorist weapon.

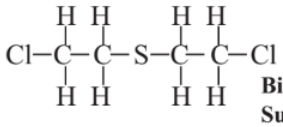
Hydrogen cyanide, HCN, is a potentially devastating gaseous pollutant. It has been used to carry out death sentences in gas chambers, causing death very rapidly when inhaled. Another toxic form of cyanide is **cyanide ion**, CN^-, in salts such as KCN. Only about 60 mg of KCN will kill a human. Glass ampules containing KCN or liquid HCN were used by some doomed Nazi leaders to commit suicide near the end of World War II. There is concern that potassium cyanide or other soluble cyanide salts may be put into water supplies as toxic agents.

The metabolic action of cyanide depends upon its strong binding for iron in the +3 oxidation state. In the essential utilization of molecular oxygen in the body (the respiration process called oxidative phosphorylation) iron cycles between iron(III) in ferricytochrome oxidase enzyme andiron(II) in the chemically reduced form, ferrous cytochrome oxidase. By stabilizing ferricytochrome oxidase, cyanide stops this cycle, preventing utilization of oxygen and causing metabolic processes to cease. It is interesting to note that an antidote to cyanide poisoning — in those rare instances where the victim survives long enough for antidotes to be administered — is to have the victim inhale a volatile nitrite ester. The nitrite converts a fraction of the iron(II) in hemoglobin to iron(III), generating methemoglobin. This form of hemoglobin cannot carry oxygen, but it can bind with cyanide, preventing it from tying up ferricytochrome oxidase enzyme.

Hydrogen sulfide, H_2S, is a colorless gas with a foul, rotten-egg odor that is as toxic as hydrogen cyanide and may kill even more rapidly. Inhalation of 1000 ppm hydrogen sulfide causes rapid death from respiratory system paralysis. Nonfatal doses of this gas can cause excitement due to damage to the central nervous system; headache and dizziness may be symptoms of exposure.

Military Poisons and Nerve Gas Agents

Beginning with the use of toxic chlorine gas in World War I, nations have developed a variety of diabolical toxic agents to disable and kill opposing troops in war. One such agent is **sulfur mustard**, a class of chemical compounds, the most common of which is bis(2-chloroethyl)sulfide



Bis(2-chloroethyl)sulfide, Sulfur Mustard

The vapors of this substance penetrate rapidly and deeply into tissue causing tissue damage and destruction well below the point of entry. Because of its penetrating ability, efforts to remove sulfur mustard from exposed tissue are futile after about 30 minutes. Sulfur mustard is classified as a vesicant ("blistering gas") producing severely inflamed lesions that are susceptible to infection. Such lesions in the lungs are likely to be fatal. (An internet search for pictures of sulfur mustard lesions will bring up some horrifying images of the wounds that this material can inflict.) Sulfur mustard causes mutations because it forms a reactive intermediate that binds with DNA and is thought to be a primary carcinogen that does not require metabolic activation to produce cancer.

The chemical agents of greatest concern for their potential use in terrorist attacks are the organophosphorus "nerve gases." The first of these deadly agents was reported in 1937 by Gerhard Schrader of the German concern Farbenfabriken Bayer AG. Work continued on these compounds in Germany during World War II and by other nations after the war and during the cold war between Western and Communist bloc nations until around 1990. The possibility that Iraq possessed large stores of military nerve gas "weapons of mass destruction" was part of the rationale for the U.S./Iraq war in 2003. Among the common nerve gases are Sarin, Soman, Tabun, CMPF, VX, and diisopropylphosphofluoridate (fluorodiisopropyl phosphate)Structural formulas of three of these compounds are shown in Figure 16.4.

Sarin is perhaps the best known organophosphorus military poison because of its use in a 1995 attack by a terrorist group on the Tokyo subway system that killed several people and caused illness in a number of others. It is estimated that a dose of only about 0.01 milligrams of Sarin per kilogram of body mass is fatal; absorption of a single drop of liquid Sarin through the skin can kill a human. Sarin and other organophosphate military poisons act on the nervous

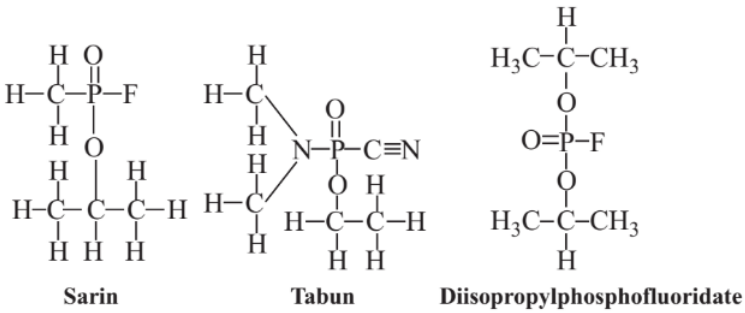
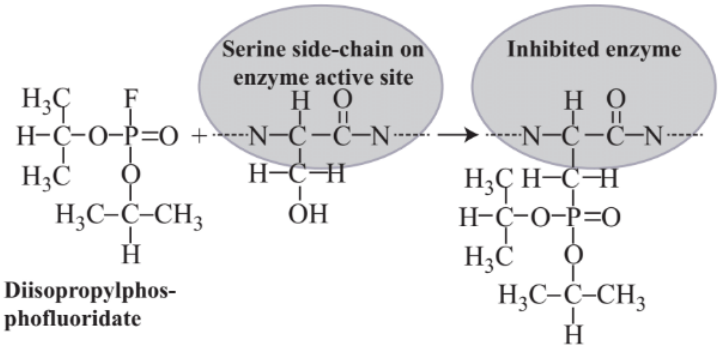


Figure 16.4. Three common nerve gas organophosphate military poisons

system by binding with and inhibiting acetylcholinesterase enzyme, which is required to hydrolyze acetylcholine and stop nerve impulses once their function has been completed. The failure of the hydrolysis of acetylcholinesterase typically causes failure of the respiratory system. The binding of diisopropylphosphofluoridate to a serine side-chain on the active site of acetylcholinesterase enzyme is shown by the reaction below:



Biotoxins

Some of the most toxic substances known are produced by organisms and some of these have been used to attack humans.*Clostridium botulinum* bacteria growing in the absence of oxygen produce **botulinum** toxin, which has killed many people who have eaten improperly canned food (heating to 80–100° C for a sufficient time deactivates the toxin). Consisting of several proteins, botulinum toxin binds irreversibly to nerve terminals preventing the release of acetylcholine, an enzyme required for proper nerve function. Neurologic symptoms are followed by paralysis of the respiratory muscles and death. From the LD50 of about 1×10^{-5} mg/kg for botulinum toxin shown in Figure 16.1, it may be inferred that a 70 kg person could die from absorbing 70 kg 1×10^{-5}mg/kg = 0.0007 mg of botulinum toxin, or only about 1 microgram of this extraordinarily deadly substance. A simple calculation would show that literally millions of people could be killed by the amount of this toxin that could be carried in a terrorist's pocket if some efficient means could be found to deliver it.

Another highly toxic natural product is **ricin**, a very stable proteinaceous material extracted from castor beans (*Ricinus communis*). Only about 1/2 milligram of ricin (about the size of a pinhead) can be fatal when injected resulting in failure of kidneys, liver, and spleen along with massive blood loss from the digestive tract. Ricin gained notoriety in the 1978 assassination in London of the Bulgarian writer and journalist Georgi Markov injected with ricin from the tip of an umbrella. Although it is mentioned as a potential terrorist tool, ricin has its greatest toxicity by injection, which tends to limit its use as a weapon.

This page titled 2.5.16.6: Protecting Water, Food, and Air is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **16.6: Protecting Water, Food, and Air** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.16.7: Detecting Hazards

In the future, disease-causing agents are more likely to be used than are chemicals for terrorist attack. Because of the ability of pathogens to multiply in the human body and to spread among people, they are much more effective than chemical agents per unit mass. A 1993 report from theU.S. Office of Technology Assessment estimated that as many as 3 million people could be killed by the airborne release of 100 kg of anthrax spores in a highly populated area such as Washington, DC. The magnitude of the collateral damage from a biological attack can be appreciated from the chaos that ensued from the October 2001 anthrax incident in which the spores for this deadly agent were spread by letters received in Washington, DC, New York, Connecticut, and Florida that resulted in 22 cases of anthrax, five of which were fatal. Congress and the U.S. Supreme Court were among the facilities closed for some time and decontamination costs came to $300 million.

In modern times the most devastating biological warfare was carried out by Japan in its biological attacks on Manchuria and China in the 1930s and during World War II. Among the pathogenic agents studied by the Japanese army were anthrax, cholera, dysentery, typhoid, paratyphoid, bubonic plague, and tularamia. In 1940 and 1941 at least three Japanese air raids against Chinese cities using bombs laden with live human fleas infected by bubonic plague killed almost 150 people. During a retreat by the Japanese army in the Chekiang Province of China in1942 water sources and houses were deliberately contaminated by cholera, plague, and dysentery agents causing many casualties among the populace and Chinese troops who occupied the area. This strategy backfired however when Japanese troops re-occupied the region suffering about 10,000 casualties of which about 1,700 were fatal due to the disease agents previously spread in the area. The release of plague agents by Japanese troop retreating from invading Soviet forces in Manchuria at the end of World War II caused plague epidemics in 1946 and 1947.

Although pathogenic organisms are the oldest form of terrorist attack, modern recombinant DNA biotechnology threatens much higher potential for devastation with these agents. The potential exists to modify known pathogens to strains that are more robust, more virulent, and resistant to antibiotics. As one particularly deadly example, Soviet scientists used recombinant DNA technology to add genes for the production of myelin toxin to *Yersiniapseudotuberculosis* bacteria so that targets infected with this pathogen would be afflicted by both tuberculosis lung symptoms and the paralysis caused by the myelin toxin.

This page titled 2.5.16.7: Detecting Hazards is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **16.7: Detecting Hazards** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.16.8: Green Chemistry to Combat Terrorism

Water to drink, food to eat, and air to breathe — the most basic of human needs — are likely targets for terrorist attack. The finding of a single cow with mad cow disease in Washington State in December, 2003, caused a major upheaval in cattle markets, loss of exports, and a final cost of perhaps hundreds of millions of dollars. Earlier outbreaks of mad cow disease and hoof and mouth disease in England had caused staggering economic loss. Mad cow disease is caused by a protein agent called a prion and can infect humans through contaminated meat causing fatal Creutzfeldt-Jakob disease characterized by devastating brain and nervous system effects. Over 200 cases ofthis malady linked to mad cow disease have been reported including a victim in Italy in 2010. Reported incidents of mad cow disease were the result of accident and poor agricultural and food production practices, not terrorism, but they illustrate the vulnerabilities of the food supply to potential terrorist attack.

A chemical attack on food supplies, though plausible, would be very difficult to carry out on a scale that would cause great damage. Spraying of food crops with toxic substances before harvest could cause some adverse effects and great anxiety, but would be relatively easy to detect and probably would not cause widespread harm.

Direct contamination of food with disease-causing agents is a possible terrorist action. An incident has been described in which 12 laboratory staff were infected by acute diarrheal illness that hospitalized 4 due to ingestion of *Shigelladysenteriae* bacteria taken from a culture in the laboratory and deliberately placed on doughnuts or muffins in the facility break room.4Mostpeople have experienced the abdominal pain, vomiting, diarrhea and fever of “food poisoning” caused by *Salmonella* bacteria in contaminated food. Hundreds of millions of eggs infected by *Salmonella* in the U.S. during the summer of 2010 caused illness in several hundred people. Although rarely fatal, *Salmonella* on food have the potential to cause temporary disability.

Because of their central distribution to large numbers of people, water supplies are susceptible to both chemical and biological attack. There are reports of terrorist groups trying to obtain deadly cyanide salts with the objective of contaminating water supplies. The tragic arsenic contamination of water from tube wells unknowingly drilled into arsenic-bearing aquifers that may have afflicted as many as 77 million people in Bangladesh is a reminder of the potential for ill effects from chemical contamination of drinking water. Astoundingly toxic botulinus toxin from *Botulinus* bacteria (see Figure 16.2) is a potential chemical agent that could be put in water supplies. Though possible, it would be rather difficult to deliberately contaminate a municipal water supply with toxic chemicals.

Throughout history, drinking water contaminated by microorganisms that cause amebic, bacterial, and viral diarrhea; cholera; typhoid; and other maladies have killed millions and still cause disease and fatalities. The World Health Organization estimates that approximately 1.8million people die each year from diarrhea and dysentery, much of which results from drinking contaminated water. In 1993, more than 400,000 people in Milwaukee were sickened and over 50 died from waterborne protozoal *Cryptosporidium parvum*. In May, 2000, approximately 3000 people were made ill and seven died in Walkerton, Ontario, Canada, from drinking water contaminated with *Escherichiacoli* bacteria. Although usually *E. coli* bacteria are harmless and normal residents of human intestinal systems, they may develop strains with DNA transferred from *Shigelladysenteriae* bacteria that produce shiga toxin, the cause of potentially fatal dysentery, which is what happened in the Ontario incident. Bacteria that could be added deliberately to drinking water include *Shigella dysenteriae*, *Vibrio cholerae*, and *Yersinia pestis.*

Air is a possible medium for both chemical and biological attack. A means is required to deliver agents through the air, which makes it difficult to expose people through this medium. Although a low-flying crop-spraying plane would be an effective means of spreading either chemical or biological agents through the air, it would rapidly alert authorities leading to corrective action. Spores of bacteria that cause anthrax, *Bacillus anthracis*, are of particular concern for bioterror attack through air. Other microbial agents of concern for their potential for airborne attack include *Variola major*, which causes smallpox;*Francisella tularensis*, which causes tularemia, and viruses that cause viral hemorrhagic fevers, including Ebola, Marburg, Lassa, and Machupo.

Historically, the disease that has caused devastation most closely resembling the harm that could result from a massive attack by bio-agents is “plague,” which killed tens of millions of people in Europe during the Middle Ages. This malady is caused by infection with *Yersiniapestis* bacteria. This disease takes several forms, the most common of which is bubonic plague characterized by swollen, tender lymph glands called buboes. Readily cured by antibiotics in its early stages, it is transferred from infected rodents to humans by fleas, and several cases are reported each year in the Southwestern U.S. Pneumonic plague is readily spread

through air between humans and is the form of most concern for terrorist attack. Initial symptoms similar to those of influenza progress to a fatal form of pneumonia.

This page titled 2.5.16.8: Green Chemistry to Combat Terrorism is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **16.8: Green Chemistry to Combat Terrorism** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.16.9: Green Chemistry for Sustainable Prosperity and a Safer World

A key to protection from terrorist threats is their detection before damage can be done. The detection of explosives immediately comes to mind. Another priority area is detection of disease-causing pathogenic organisms.

Hazardous substances are not readily detected by standard metal detectors and X-ray imagers used to find weapons and bombs on air travelers and in their luggage. Residues of TNT, RDX, and PETN explosives (see Figure 16.1) can be detected by sophisticated instruments including ion mobility spectrometers and chemiluminescence sensors. Such instruments normally detect residues of explosives on swabs from swabbing luggage; they can be circumvented by careful cleaning of luggage. Nuclear quadrupole resonance (NQR) may develop as an especially promising detection technique for explosives. One reason for this is its specificity for nitrogen, which is abundantly present in all common explosives. Secondly, NQR has the potential to detect explosives in containers and even land mines. It works by generating a pulse of radio frequency radiation which excites nitrogen atoms to higher quantized energy levels. By following the signals given off as the atoms return to their ground energy levels the kinds and abundances of nitrogen functional groups in explosives molecules can be determined.

“Canine olfactory detection,” which uses the sniffing abilities of dogs, is widely used to detect explosives, illicit drugs, and other potentially hazardous materials. A dog has approximately 220 million mucus-coated olfactory receptors, about 40 times the number possessed by a human, making the canine nose an extraordinarily sensitive detector. In order for canine olfactory detection to work, a rewards system must be used. This can lead to unpredictable, temperamental behavior in dogs of the type commonly attributed to humans and computers. As a result, dogs are not completely reliable and, according to an authority on the subject, “Dogs lie. We know they do.”[5]

This page titled 2.5.16.9: Green Chemistry for Sustainable Prosperity and a Safer World is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **16.9: Green Chemistry for Sustainable Prosperity and a Safer World** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.16.10: Green Chemistry to Combat Terrorism

As *safe* and *sustainable* chemistry, green chemistry has an important role to play in the war on terrorism. When safe chemistry is practiced, hazards and hazardous substances that might be stolen or diverted for use in attacks are not made or used in large quantities. The successful practice of green chemistry means that chemical products do what they are supposed to do and are used in minimum quantities. With green chemistry, materials and processes that are likely to result in violent reactions, fires, high pressures, and other extreme conditions are avoided and auxiliary substances and flammable materials are not used.

Green chemistry minimizes energy consumption, thereby reducing energetic, high-temperature processes that might be susceptible to sabotage. With its emphasis on biological processes, where applicable, green chemical processes are carried out under the mild, low temperature, toxic-substance-free, inherently less hazardous conditions conducive to biochemical reactions. By reducing demand for energy and raw materials, green chemistry reduces reliance on uncertain sources that are controlled by potentially hostile populations and therefore are inherently subject to disruption and blackmail.

The practice of green chemistry requires exacting process control combined with real-time,in-process monitoring techniques. These are conditions that resist sabotage. In addition, green chemistry uses passive safety systems that function by default in the event of failure of or deliberate damage to intricate control systems. An example of such a system is gravity-fed cooling water for nuclear reactors that will continue to flow even if the pumps for the cooling system pumps fail.

The chemical industry and related enterprises continue to implement green chemistry practices to reduce hazards and vulnerabilities to attack. For example, a DuPont chemical facility in Texas now makes methyl isocyanate, the agent of the catastrophic 1984 Bhopal, India, disaster, on site so that it does not have to store large quantities of this dangerous chemical.

As noted in Section 16.6, liquified and pressurized elemental chlorine, the first “poison gas” to be employed in warfare, has a high potential for terrorist attack because of its toxicity, volatility, and widespread shipping and storage. Alkaline aqueous solutions of hypochlorite ion, OCl^-, are widely used for bleaching and disinfection. Because liquified Cl_2 is the most economical form in which to ship chlorine, the common practice has been to react the elemental substance with base to produce hypochlorite:

$$Cl_2 + 2\,Na^+ + 2\,OH^- \mhchemleftrightarrows 2\,Na^+ + OCl^- + Cl^- + H_2O$$

Maintained in a basic state, the equilibrium of this reaction lies to the right so that the equilibrium concentration of Cl_2 in solution remains very low. Both Cl_2 and NaOH are made by passing a direct electrical current through a solution of NaCl where the reaction at the cathode is

$$2\,Na^+ + 2\,H_2O + 2\,e^- \mhchemleftrightarrows 2\,Na^+ + 2\,OH^- + H_2\,(g)$$

and that at the anode is the following (where e^- represents the electron)

$$2\,Cl^- \mhchemleftrightarrows Cl_2\,(g) + 2\,e^-$$

The chlorine gas produced at the anode can be reacted with the basic NaOH solution generated at the cathode to produce sodium hypochlorite (Reaction 16.8.1), which is produced, stored, and used as aqueous solutions. Many water treatment plants have started using relatively safe sodium hypochlorite solutions in place of toxic, reactive liquid chlorine formerly stored in large pressurized tanks on site. A potential problem in storing sodium hypochlorite solutions is their decomposition to produce elemental oxygen,

$$2\,Na^+ + 2\,OCl^- \rightarrow O_2\,(g) + 2\,Na^+ + 2\,Cl^-$$

a process catalyzed by traces of transition metal ions such as Cu^{2+}. This results in a loss of product and a potentially hazardous buildup of oxygen and pressure.

By 2010 the Clorox company, a major world supplier of chlorine-based bleach announced that it would no longer be using liquid chlorine to make sodium hypochlorite, but would be shipping relatively concentrated aqueous solutions of sodium hypochlorite and diluting them to the strength required for bleaching. A reasonable alternative to using liquid chlorine or solutions of sodium hypochlorite for larger installations, such as major metropolitan water treatment plants, would be to have relatively small installations for the electrolysis of sodium chloride solutions which would make sodium hypochlorite directly on site. Such a facility might fit well with an industrial ecosystem in which several users of sodium hypochlorite would be clustered in close proximity.

2.5.16.10: Green Chemistry to Combat Terrorism is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **16.10: Green Chemistry to Combat Terrorism** has no license indicated.

2.5.16.11: Green Chemistry for Sustainable Prosperity and and a Safer World

Poverty, human misery, and hopelessness are conditions that feed terrorism. Although eliminating these conditions would not guarantee a safe world, reducing them would go far toward ensuring safe societies. People with satisfied material needs able to lead comfortable and fulfilling lives are relatively less likely to commit violent acts. To the extent that the practice of green chemistry fulfills human needs and makes life more comfortable, it can play a significant role in reducing terrorism.

Prosperity, narrowly defined, has resulted in consumption of increasingly scarce resources and environmental degradation. But, as stated by Elsa Reichmanis, a former President of the American Chemical Society, “We are past the days when we can trade environmental contamination for economic prosperity; that is only a temporary bargain, and the cost of pollution both economically and on human health is too high.”[6] Green chemistry and the practice of industrial ecology can go far in providing high living standards without ruining the environment or recklessly exploiting resources.

The key to material prosperity consists of sources of abundant, inexpensive energy that can be tapped sustainably without major environmental harm; with such energy sources, all else is possible. Energy sources tend to be contentious and competition for them has precipitated past wars. Some of the most abundant producers of petroleum, currently the key energy source for industrialized nations, are regions that are breeding grounds for terrorists. The provision of adequate energy independent of such sources would substantially reduce terrorist threats.

Figure 16.5 shows how abundant, sustainable energy is the foundation of the kind of prosperity that can lead to less terror-prone societies. Abundant energy can be used to produce food through synthesis of fertilizers (particularly by synthetic fixation of atmospheric nitrogen)and for irrigation, cultivation, and reclamation of farmland. Energy can be used to fabricate materials for housing and to provide the heating, cooling, and lighting required to make dwellings comfortable. Energy is required to pump water, in some cases over great distances from abundant sources to more arid regions. Energy can be used to purify water of marginal quality and to reclaim water after use. With an abundant source of energy, seawater can be desalinated for domestic, industrial, and agricultural use. Safe, comfortable, non-polluting transportation systems require an abundance of energy. These and other amenities based upon abundant, sustainable energy can go far toward building peaceful, productive societies with high living standards. They do not guarantee tranquility and prosperity because reasonable social systems, functional democratic governments, and sensible religions are needed as well, but material well being based upon a foundation of abundant, sustainable energy can go far in eliminating conditions that breed terrorism.

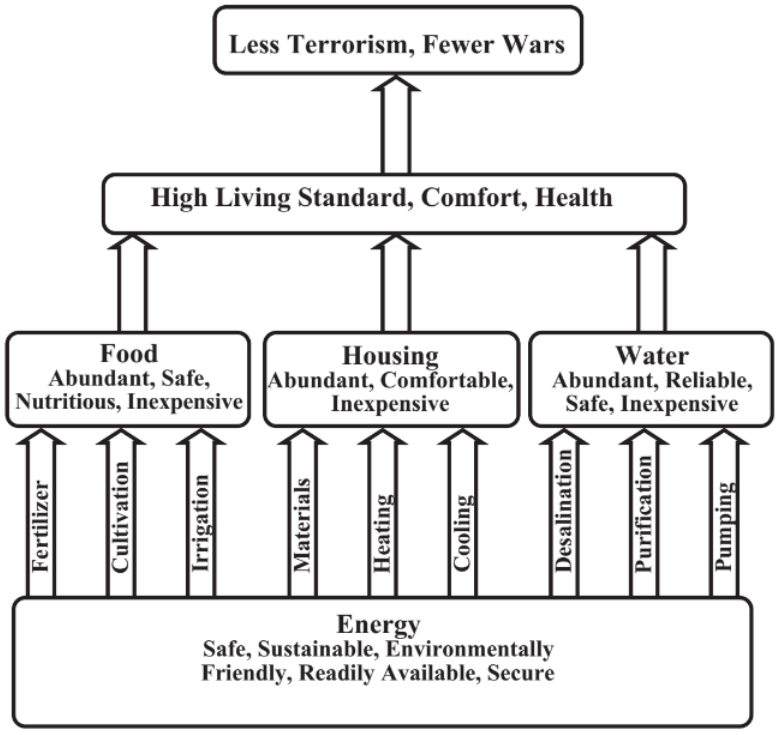


Figure 16.5. Abundant, sustainable energy is the base of a pyramid through which greater human well-being combined with suitable political and social systems leads to less conflict and terrorism

The provision of abundant, sustainable energy in the future requires the best practice of green chemistry, green engineering, and industrial ecology. Increased efficiency of energy utilization is a key aspect of providing more usable energy. Solar, wind, and biomass energy are leading contenders for renewable energy sources. Another essentially inexhaustible energy source is thermonuclear fusion, the stuff of hydrogen bombs and the sun's energy, but despite significant investments, a practical controlled system of energy production from this source has proven elusive. Fossil fuels will play an interim role, especially if sequestration of greenhouse gas carbon dioxide byproduct can be achieved. Despite its bad reputation in some quarters, nuclear fission with uranium and perhaps thorium fuel can provide abundant energy safely with new-generation nuclear reactors and with reprocessing of nuclear fuel.

A key challenge in providing abundant renewable energy is its storage and transport. Wind and solar sources are by nature intermittent and dispersed, and they often produce electricity in locations far from where it is used, so the energy that they generate must be moved over long distances and stored for later use. For example, solar collectors function only in daytime and, aside from rooftop installations, are often located in remote desert locations. Wind-powered electrical generators, which require at least some wind, are not usually welcome in urban areas where the energy is required, and some of the prime locations for them are the remote plains of Kansas orTexas or offshore. Superconductor or quantum conductor power cables are candidates for transport of electrical energy from source to use. Various means are available for energy storage, such as pumped water hydroelectric storage or high-speed flywheels coupled with electric motor/generators. In the future, elemental hydrogen, H2, will be widely used for energy storage and transport as well as for fuel. Hydrogen can be produced by electrolysis of water and direct photoconversion of water to hydrogen and oxygen may eventually become practical. Hydrogen can be moved by pipeline and used to produce electricity directly in fuel cells.

2.5.16.11: Green Chemistry for Sustainable Prosperity and and a Safer World is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **16.11: Green Chemistry for Sustainable Prosperity and and a Safer World** has no license indicated.

2.5.16.12: Literature Cited

1. Sachs, Jeffrey D., “Blackouts and Cascading Failures of the Global Markets,” *Scientific American*, January, 2009.

2. Biesecker, Calvin, “Report Warns of Cyber Threat to U.S. Supply Chains,” *Defense Daily* 247,August 5, 2010.

3. Spiers, Edward M., *A History of Biological and Chemical Weapons*, University of Chicago Press, Chicago, 2010.

4. Kolavic, Shellie A., Akiko Kimura, Shauna L. Simons, Laurence Slutsker, Suzanne Barth, and Charles E.*Haley*, “An Outbreak of *Shigelladysenteriae* Type 2 Among Laboratory Workers Due to Intentional Food Contamination,” *Journal of the American Medical Association*,278,396–8 (1997).

5. Derr, Mark, “With Dog Detectives, Mistakes Can Happen,” *New York Times*, Dec. 24, 2002, p.D1.

6. As quoted in “Green Solutions to Global *Problems*,” Stephen K. Ritter, *Chemical and Engineering News*,81, 31–33 (2003).

2.5.16.12: Literature Cited is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Literature Cited** has no license indicated.

2.5.16.13: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right."

1. At the beginning of this chapter, it was noted that the Alfred P. Murrah Oklahoma City Federal Building was destroyed in 1995 by the explosion of a mixture of ammonium nitrate (chemical formula NH_4NO_3) and diesel fuel ($C_{16}H_{34}$). An explosion occurs when chemicals react very rapidly to produce much more stable chemical species, gases, and heat. Consider that H_2O, CO_2, and N_2 are very stable forms of the elements present in a mixture of ammonium nitrate and diesel fuel, that CO_2, and N_2 are gases, and that at the temperature of an explosion H_2O is a gas as well. Attempt to write a chemical reaction that occurs when a mixture of ammonium nitrate and diesel fuel explodes.

2. Nitroglycerin explodes by itself without having to react with any other chemical. Look up its formula and, from the information given in the preceding question, give a possible chemical reaction for a nitroglycerin explosion.

3. Consider the infrastructure of your home. Suggest how it might be vulnerable because of the interconnectivity in it. Suggest how cascading failures might do great damage to your homelife.

4. For a 75-kg person, estimate the lethal dose in grams, milligrams, or micrograms of each of the toxic substances shown in Figure 16.2.

5. Exposure of a person to toxic benzene can be estimated by measuring phenol in blood. Explain the rationale for such an analysis. Why is benzene epoxide not commonly determined to estimate benzene exposure?

6. Consider the toxicity of inhaled carbon monoxide in the context of Figure 16.3. Identify for carbon monoxide the receptor, the abnormal biochemical effect, and the physiological response manifesting toxicity.

7. Compare carbon monoxide to chlorine as agents of terrorist attack. What characteristic of carbon monoxide might make it a favored weapon. Why might chlorine be favored by terrorists? What characteristic would make it less attractive than carbon monoxide to terrorists?

8. Match each toxic substance and potential terrorist agent from the list on the left below with its characteristic from the list on the right.

A. Botulinus toxin 1. A disease-causing agent (pathogen)

B. Sarin 2. Inhibits acetylcholinesterase enzyme

C. Methyl isocyanate 3. Most toxic of those listed

D. Shigella dysenteriae 4. Toxic in gaseous, solid, or solution form

E. Cyanide 5. An industrial chemical that has killed thousands in a single incident

9. Figure 16.5 illustrating how abundant, sustainable sources of energy can lead to a high living standard emphasizes food, housing, and water. Suggest areas other than these three that depend upon abundant, sustainable energy and that can lead to a high living standard.

10. Natural gas pipelines are sometimes cited as elements of the infrastructure that are vulnerable targets for terrorist attack. Look up and describe an incident from 2010 that illustrates this vulnerability.

11. The compound sodium chlorate, $NaClO_3$, is mentioned in this chapter as an oxidant. Suggest with a chemical reaction what happens to this compound when it is heated. How might it actas an agent to accelerate fire or even cause an explosion?

12. What is meant by "just-in-time" chemical production and how may it minimize hazards in chemical production? In what sense is it consistent with the practice of green chemistry? In what sense does it potentially contribute to vulnerability in the chemical manufacture process?

13. Native American populations in what is now the U.S. were drastically reduced by a disease-causing agent after European explorers and settlers arrived on the continent. What was this agent and why were the Native Americans particularly vulnerable to its effects? Were there any cases in which it was deliberately spread?

14. There are two major ways in which biological agents (including microorganisms) can be used to harm humans. Explain.

15. At least two different water disinfecting agents can be made with electricity using raw materials so abundant that they may be regarded as renewable. What are these agents and how could they be synthesized as needed on site at a water treatment plant? How might such systems reduce the danger of terrorist attack?

2.5.16.13: Questions and Problems is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **Questions and Problems** has no license indicated.

SECTION OVERVIEW

2.5.17: The Ten Commandments of Sustainability and Sensible Measure

“On August 14, 2010, a massive traffic jam began building on one of the main highways into Bejing, China, that would eventually extend for 60 miles and last for weeks, on some days moving less than a mile. Nothing better illustrates the futility of continued reliance on individual vehicles for basic transportation needs in an increasingly crowded world.”

2.5.17.1: We Cannot Go on Like This

2.5.17.2: The First Commandment

2.5.17.3: The Second Commandment

2.5.17.4: The Third Commandment

2.5.17.5: The Fourth Commandment

2.5.17.6: The Fifth Commandment

2.5.17.7: The Sixth Commandment

2.5.17.8: The Seventh Commandment

2.5.17.9: The Ninth Commandment

2.5.17.10: The Tenth Commandment

2.5.17.11: Some Sensible Measures for Sustainability

2.5.17.12: The Eighth Commandment

2.5.17.13: Literature Cited

2.5.17.14: Questions and Problems

2.5.17.15: Supplementary References

This page titled 2.5.17: The Ten Commandments of Sustainability and Sensible Measure is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.17.1: We Cannot Go on Like This

In 1968 the Stanford University biologist Paul Ehrlich published a book entitled *The Population Bomb*,[1] a pessimistic work that warned Earth had reached its population carrying capacity sometime in the past and that catastrophe loomed. Ehrlich predicted rapid resource depletion, species extinction, grinding poverty, starvation, and a massive dying of human populations in the relatively near future. “Not so,” retorted Julian Simon a University of Maryland economist writing in a number of books, the most recent of which, finished just before his death in 1998, is titled *Hoodwinking the Nation*.[2] Ehrlich hedged his views by stating that he might be wrong and that “some miraculous change in human behavior” or a “totally unanticipated miracle” might “save the day.” Simon expressed the view that Ehrlich’s doom and gloom views were nonsense and that human ingenuity would overcome the problems foreseen by him.

The debate between Ehrlich and Simon led to a famous wager by Simon in 1980 that $200 worth of each of five raw materials chosen by Ehrlich—copper, chromium, nickel, tin and tungsten—would actually decrease in price over the next 10 years in 1980 dollars. Each did in fact decrease in price and Ehrlich paid. Simon then offered to raise the ante to $20,000, a proposition that Ehrlich declined. (After 1990 there were some spectacular price increases in these and other mineral commodities.

The Ehrlich/Simon wager is often cited by anti-environmentalists as evidence that we will never run out of essential resources and that a way will always be found to overcome shortages. However, common sense dictates that Earth’s resources are finite. Whereas unexpected discoveries, ingenious methods for extracting resources, and uses of substitute materials will certainly extend resources, a point will inevitably be reached at which no more remains and modern civilization will be in real trouble.

Unfortunately, the conventional economic view of resources often fails to consider the environmental harm done in exploiting additional resources. Fossil fuels provide an excellent example. As of 2005, there was ample evidence that world petroleum resources were strained as prices for petroleum reached painfully high levels. This has resulted in a flurry of exploration activities including even drilling in some cemeteries! Natural gas supplies have been extended by measures such as tapping coal seams for their gas content, often requiring pumping of large quantities of alkaline water from the seams and release of the polluted water to surface waters. There is no doubt that liquid and gaseous fossil fuel supplies could be extended by decades using coal liquefaction and gasification and extraction of liquid hydrocarbons from oil shale. But these measures would cause major environmental disruption from coal mining and processing, production of salt-laden oil shale ash, and release of greenhouse gases.

The sad fact is that on its present course humankind will deplete Earth’s resources and damage its environment to an extent that conditions for human existence on the planet will be seriously compromised or even become impossible. There is ample evidence that in the past civilizations have declined and entire populations have died out because key environmental support systems were degraded.3A commonly cited example is that of the Easter Islands where civilizations once thrived and the people erected massive stone statues that stand today. The populations of these islands vanished and it is surmised that the cause was denuding the islands of once abundant forests. A similar thing happened to pre-Columbian Viking civilizations in Greenland, where 3 centuries of unusually cold weather and the Vikings’ refusal to adopt the ways of their resourceful Inuit neighbors were contributing factors to their demise. Iceland almost suffered a similar fate, but the people learned to preserve their support systems so that Iceland is now a viable country.

There is much truth in the expression that, “The only thing we have to do to ensure a planet uninhabitable by our descendants is to continue what we are doing now.”Fortunately, modern civilizations have the capacity to avoid the fates of the ancient Easter Islanders and GreenlandVikings—if they can muster the will and the institutional framework to do so. The key is sustainability, which simply means living in ways that do not deplete Earth’s vital support systems. The great challenges to sustainability are (1) population growth beyond Earth’s carrying capacity, (2) potentially disruptive changes in global climate, (3) provision of adequate food, (4)depletion of Earth’s resources, (5) supply of adequate energy, and (6) contamination of Earth’senvironment with toxic and persistent substances. It won’t be easy to overcome these challenges and achieve sustainability and it is by no means certain that humankind will ultimately succeed or even survive on Earth. But we have to try; the alternative of a world population reduced to just a few million people surviving in poverty and misery on a sadly depleted planet under conditions hostile to higher life forms is too grim to contemplate.

The achievement of sustainability will require adherence to some important principles. These can be condensed into ten commandments of sustainability, which are listed below:

1. Human welfare must be measured in terms of quality of life, not just acquisition of material possessions, which demands that economics, governmental systems, creeds,and personal life-styles must consider environment and sustainability.

2. Since the burden upon Earth's support system is given by the relationship Burden = (number of people)×(demand per person) it is essential to address both numbers of people on Earth and the demand that each puts on Earth's resources.

3. Given that even at the risk of global catastrophe, *technology will be used* in attempts to meet human needs, it is essential to acknowledge the anthrosphere as one of the five basic spheres of the environment and to design and operate it with a goal of zero environmental impact and maximum sustainability.

4. Given that energy is a key to sustainability, the development of efficiently-used, abundant sources of energy that have little or no environmental impact is essential

5. Climate conducive to life on Earth must be maintained and acceptable means must be found to deal with climate changes that inevitably occur.

6. Earth's capacity for biological and food productivity must be maintained and enhanced, considering all five environmental spheres.

7. Material demand must be drastically reduced; materials must come from renewable sources, be recyclable and, if discarded to the environment, be degradable

8. The production and use of toxic, dangerous, persistent substances should be minimized and such substances should not be released to the environment; any wastes disposed to disposal sites should be converted to nonhazardous forms.

9. It must be acknowledged that there are risks in taking no risks.

10. Education in sustainability is essential; it must extend to all ages and strata of society, it must be promulgated through all media, and it is the responsibility of all who have expertise in sustainability.

Each of the ten commandments of sustainability is discussed in the remaining sections of this Chapter.

This page titled 2.5.17.1: We Cannot Go on Like This is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **17.1: We Cannot Go on Like This** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.17.2: The First Commandment

HUMAN WELFARE MUST BE MEASURED IN TERMS OF QUALITY OF LIFE, NOT JUST ACQUISITION OF MATERIAL POSSESSIONS, WHICH DEMANDS THAT ECONOMICS, GOVERNMENTAL SYSTEMS, CREEDS, AND PERSONAL LIFE-STYLES MUST CONSIDER ENVIRONMENT AND SUSTAINABILITY.

This commandment goes to the core question of, "What is happiness?" Many people have come to measure their happiness in terms of material possessions—the larger sport utility vehicle, the bigger house on a more spacious lot farther from the city, more and richer food. But such measures of human welfare based upon the accumulation of more stuff has come at a high cost to Earth as a whole and even to the people who acquire the most stuff. The sport utility vehicle guzzles fuel from steadily decreasing petroleum supplies, commodious houses require more energy to heat and cool, large lots remove increasingly scarce farm land from food production, dwellings far from the workplace mean long commutes that consume time and fuel, and too many of the current generation of humans have consumed food to a state of unhealthy obesity.

The things that really count for happiness—good health, good nutrition, physical comfort, satisfying jobs, good interpersonal relations, interesting cultural activities—can be had with much less consumption of materials and energy than is now the case in wealthier societies. In order for sustainability to be achieved, it is essential for societies to recognize that happiness and well-being are possible with much less consumption of materials and energy.

Environmental and Sustainability Economics

Conventional neoclassical (Newtonian) market economics have not adequately considered resource and environmental factors in the overall scheme of economics. Since about 1970, however, environmental and natural resource economics has developed as a viable discipline.[4] This discipline, commonly called environmental economics addresses the failure of a strictly market economy to deal with scarcity and to address environmental problems. Much more complex than neoclassical economics, environmental economics addresses sustainability issues, resource economics, pollution costs, costs and benefits of pollution control, and the value of natural capital. Economic instruments can be powerful influences in reducing pollution and extending resources. The conventional market economy does act to extend resources. For example, as petroleum prices increase to painfully high levels, the rate at which consumption increases is diminished. Artificial market intervention can act to thwart such a desirable outcome. For example, U.S. government subsidies of biomass-based ethanol and biodiesel fuels contribute to increased stress on agricultural resources requiring increased amounts of fertilizers and fuel to grow the extra grain required to produce grain-based fuels.

Economic measures can be used to reduce pollution and demand on resources. Carbon and energy taxes can be imposed to reduce emissions of greenhouse gas carbon. Pollution trading has evolved as an effective pollution control measure. In the case of greenhouse gas carbon dioxide, for example, a utility installing a new coal-fired power plant may pay another concern to do reforestation projects that take an equivalent amount of carbon dioxide from the atmosphere.

More difficult to quantify, but no less real, are environmental amenities. There are certainly costs associated with impaired air quality in terms of increased respiratory disease and damage to buildings. In principle, such costs are quantifiable. Much more difficult to quantify are the value of a beautiful scenic view or the costs of eyesore billboard clutter.

A major issue with environmental economics is that of expenditures in the public sector versus those in the private sector. Free market capitalism is a powerful force in providing goods and services and in promoting innovation. Dismal past failures of planned economies and subsequent growth of these economies after they were converted to free market systems—China is probably the most striking example—illustrate the power of market forces. However, much of what is needed for sustainable development requires investment in the public sector, especially in infrastructure. The central challenge for economic systems in the future will be to integrate essential development in the public sector with free market forces. Both are essential in order for sustainable economic systems to flourish.

The Role of Governments

Sustainability will require the strong involvement of governments at all levels and extending across international boundaries. At local levels ordinances and regulations that promote sustainability are essential. For example, there are many cases in which local governments have set up recycling programs for paper, plastic, glass, and metals to reduce the need to dispose of solid wastes. In many cases only national governments have the power and authority to undertake massive projects and to promote changes

required for sustainability. Since sustainability is a global concern, ways must be found to enable governmental action and cooperation among nations.

An essential part of the role of government in sustainability is the quality of government and the people involved in it as well as the public perception of government. “Government bashing” is fashionable in many circles, and in some cases is even richly deserved. However, in order for sustainability to succeed, the finest minds that societies have must be willing to enter government service and their contributions must be respected by the public.

Personal Life Styles and Value Systems

The achievement of sustainability will require an unprecedented commitment from individuals. This may well be the most difficult of all objectives to achieve. Many people seem to have an insatiable appetite for possessions and activities that consume large amounts of materials and energy. Nothing illustrates this better than the private automobile; most teenagers find the wait to get their driver’s licenses excruciatingly long and senior citizens dread the day when they are no longer able to drive.

Although people in developed countries are commonly accused of being too materialistic, populations in less developed countries have the same desires for material possessions. Some of the greatest environmental and resource impacts occur when the economies of less developed nations improve to the point that large numbers of their citizens can afford more of the things and services that prosperity, conventionally defined, offers. For example, as of 2010, the fastestgrowing market for automobiles was in China as its economy grew.

The achievement of sustainability will require that individuals adopt sustainability as part of their belief systems. Indeed, it would be very helpful if environmental protection and the preservation of Mother Earth and her limited resources were to become virtually a religion or to be incorporated into existing religions. In this respect, some of the more primitive of Earth’s tribes had belief systems that were much more consistent with sustainability than the predominant religions of today. In some pre-Columbian Native American cultures, Earth and nature were worshipped, a belief system that could well serve as an example to current denizens of the globe. There is some evidence that modern religions are beginning to consider sustainability as a moral issue. One example is the movement, “What would Jesus drive?,” that preaches that pollution from vehicles significantly impacts human health, peace and security are threatened by reliance on imported oil from politically unstable regions, and, therefore, Jesus would not likely drive a fuel guzzling sport utility vehicle!

This page titled 2.5.17.2: The First Commandment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **17.2: The First Commandment** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.17.3: The Second Commandment

GIVEN THAT THE BURDEN UPON EARTH'SSUPPORT SYSTEM IS THE PRODUCT OF NUMBER OF PEOPLE TIMES DEMAND PER PERSON, IT IS ESSENTIAL TO ADDRESS BOTH NUMBERS OF PEOPLE ON EARTH AND THE DEMAND THAT EACH PUTS ON EARTH'S RESOURCES

The burden placed upon Earth's support systems can be expressed by the equation

$$\text{Burden} = (\text{Number of people}) \times (\text{Demand per person})$$

This equation shows that both the number of people and the demand that each puts on Earth'sresources must be considered in reducing the impact of humans on Earth. Both must be addressed to achieve sustainability.

As of 2010, Earth's human population approached 7 billion people and that of the U.S. stood at approximately 308 million people. Numbers of people on Earth will be controlled eventually by one way or another. Studies of natural ecosystems show that increases in population beyond an ecosystem's carrying capacity are always followed by a population crash; the same principle applies to human populations. Although improved food crop productivity and other measures have so far averted the catastrophic starvation that some experts were predicting in the 1960s, these humane practices have simply delayed the massive population crash that inevitably will occur if population continues to grow.

There is some good news (or perhaps less bad news) regarding population in that the population increase is significantly below that suggested by projections made 40 or 50 years earlier. Even in developing countries, birth rates have fallen to significantly lower levels than expected earlier. Particularly in Italy, Spain, France, and other nations in Europe, birth rates have fallen to appreciably below the replacement level and there is concern over depopulation. Even in the U.S., the birth rate has fallen below replacement levels and population growth that is taking place is the result of immigration. The increase in world population that has occurred over the last half century has been more due to decreasing death rates than to increasing birth rates. One U.N.official opined that, "It is not so much that people started reproducing like rabbits that they stopped dying like flies!" Although these trends do not justify complacency—population growth rates globally and in the world as a whole are still too high—they provide room for some encouragement and give hope that the first factor in Equation 17.3.1 may be controlled.

Reduction of demand per person for materials, energy, and other natural capital supplied by Earth through the application of sustainability science and technology must be accomplished to help mitigate the demands of high population. The second factor in the above equation, demand per person, may prove to be more intractable. A reasonable indicator of demand is reflected in the amount of carbon emitted per person each year, which reflects fossil fuel consumption as shown for several major countries in Figure 17.1. This figure shows that the more industrially developed countries emit the most per capita. However, the two countries with the largest populations, Chinaand India, have much lower carbon emissions per person. As the economies of these two giants grow, demand for material goods and energy-consuming services will grow as well. For example,if the average living standard of the citizens of China were to reach the modest level of those of Mexico, world petroleum consumption would have to double under conventional economic systems. Were the average person in China to live like the average person in the U.S., an impossible burden would be placed on Earth's carrying capacity. Obviously, ways must be found to meet the basic resource demands per person in more developed countries and means found to deliver a high quality of life to residents of less developed countries without placing unsupportable demands on Earth's resources.

Figure 17.1 illustrates another point regarding the relationship of population and consumption per capita, that the addition of population to more developed countries has a much greater impact on resources than it does on less developed nations. Taking per capita carbon emissions as a measure of impact, the addition of one person to the U.S. population has *at least 10 times* the impact as adding one person to India's population. It may be inferred that immigration into the U.S. and other developed countries from less highly developed nations has an inordinate impact upon resources as the immigrants attain the living standards of their new countries.

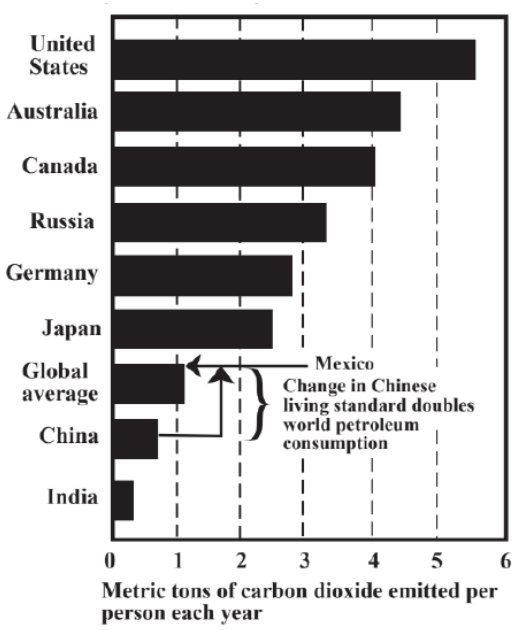


Figure 17.1. Per capita emissions of carbon dioxide per year in some major countries. The data shown are from 2004.

This page titled 2.5.17.3: The Second Commandment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **17.3: The Second Commandment** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.17.4: The Third Commandment

GIVEN THAT EVEN AT THE RISK OF GLOBAL CATASTROPHE, ***TECHNOLOGY WILL BE USED*** IN ATTEMPTS TO MEET HUMAN NEEDS, IT IS ESSENTIAL TO ACKNOWLEDGE THE ANTHROSPHERE AS ONE OF THE FIVE BASIC SPHERES OF THE ENVIRONMENT AND TO DESIGN AND OPERATE IT WITH A GOAL OF ZERO ENVIRONMENTAL IMPACT AND MAXIMUM SUSTAINABILITY

One of the most counterproductive attitudes of some environmentalists is a hostility to technology and to technological solutions to environmental problems. Humans are simply not going to go back to living in caves and teepees. Technology is here to stay. And even recognizing that the misuse of technology could result in catastrophe, *it will be used* to attempt to fulfill human needs. To deny that is unrealistic and foolish.

So a challenge for modern humankind is to use technology in ways that do not irreparably damage the environment and deplete Earth's resources. In so doing it is essential to recognize the anthrosphere—structures and systems in the environment designed, constructed, and modified by humans—as one of the five main spheres of the environment (see Chapter 8, Section 8.6). Some of the major parts of the anthrosphere are shown in Figure 17.2

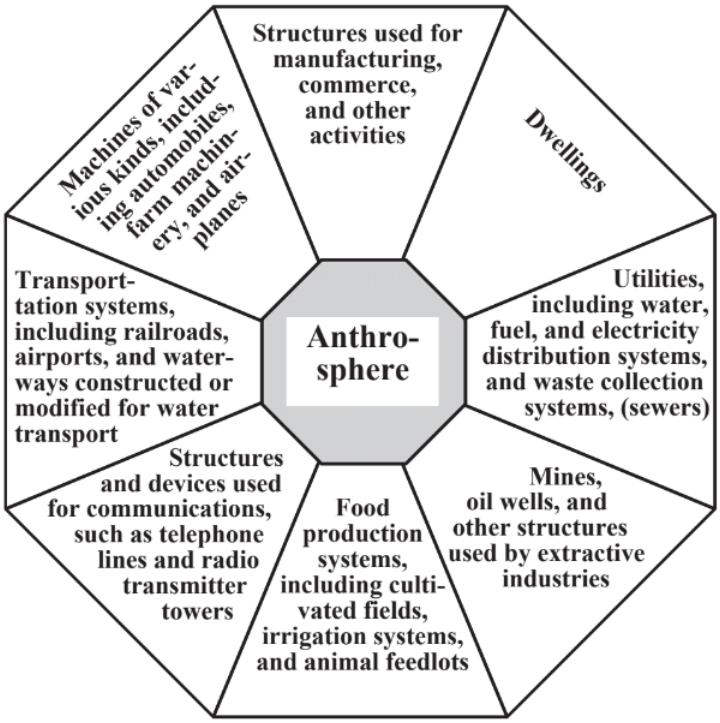


Figure 17.2. The anthrosphere is a multifaceted system designed and constructed by humans.

A key to sustainability is reorientation of the anthrosphere so that (1) it does not detract from sustainability and (2) it makes a contribution to sustainability. There is enormous potential for improvement in both of these areas.

Much is already known about designing and operating the anthrosphere so that it does not detract from sustainability. This goal can be accomplished through applications of the principles of industrial ecology discussed in Chapter 13. Basically, the anthrosphere must be operated so that maximum recycling of materials occurs, the least possible amount of wastes are generated, the environment is not polluted, and energy is used most efficiently. Furthermore, to the maximum extent possible, materials and energy must come from renewable sources.

The anthrosphere can be designed and operated in a positive way to improve and enhance the other environmental spheres. For example, modern earth-moving equipment with its capacity to move enormous amounts of material, though largely used in the past in ways that harmed the environment, can be employed positively to modify the geosphere surface in ways that will enhance the biosphere. Some specific examples of things that can be done are the following:

- Restoration of topsoil in areas depleted of this resource by poor farming practices or by contamination by wastes and pollutants

- Terracing land to prevent water erosion of soil
- Removal of obsolete and abandoned anthrospheric structures, such as old steel mills, and decontamination and restoration of the sites upon which they are located
- Construction of wetlands that can serve to restore wastewater to a quality enabling its release to the environment
- Pumping water underground to restore depleted aquifers
- Addition of “meanders” to streams, some of which have been unwisely straightened in the past, to reduce erosion and flooding
- Dredging of sediments from bodies of water and sediments to restore conditions conducive to aquatic life
- Construction and operation of reverse osmosis plants to remove excess salt from irrigation waters
- Construction of electrified railroads to replace inefficient, resource intensive, environmentally damaging truck transport

---

This page titled 2.5.17.4: The Third Commandment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **17.4: The Third Commandment** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.17.5: The Fourth Commandment

GIVEN THAT ENERGY IS A KEY TO SUSTAINABILITY, THE DEVELOPMENT OF EFFICIENTLY-USED, ABUNDANT SOURCES OF ENERGY THAT HAVE LITTLE OR NO ENVIRONMENTAL IMPACT IS ESSENTIAL

As discussed in Chapter 15, “Sustainable Energy: The Essential Basis of Green Systems,” abundant, sustainable energy that can be used without harming the environment is arguably the most important facet of sustainability. Several aspects of energy in sustainability are discussed here. Examples of what can be accomplished with sufficient sustainable energy include the following: Toxic organic matter in hazardous waste substances can be totally destroyed and any remaining elements can be reclaimed or put into a form in which they cannot pose any hazards, wastewater from sewage can be purified to a form in which it can be reused as drinking water, pollutants can be removed from stack gas, and essential infrastructure can be constructed.

The accomplishment of sustainability is impossible without the development of efficient, sustainable, nonpolluting sources of energy. Here lies the greatest challenge to sustainability because the major fossil-fuel-based energy sources used today are inefficient, unsustainable, and, because of the threat to world climate from greenhouse gases, threaten Earth with a devastating form of pollution. Other means to provide energy that are friendly to the environment and sustainable must be developed.

Fortunately, as discussed in Chapter 15, alternatives are available to fossil fuels, given the will to develop them. Most renewable energy sources are powered ultimately by the sun. The most direct use of solar energy is solar heating. Solar heating of buildings and of water has been practiced increasingly in recent decades and should be employed wherever possible. The conversion of solar energy to electrical energy with photovoltaic cells is feasible and also practiced on an increasing scale. At present, electricity from this source is more expensive than that from fossil fuel sources, but solar electricity is gradually coming down in price and is already competitive in some remote locations far from power distribution grids.

Wind energy has emerged as a somewhat surprising alternative to fossil fuels and is now competitive in price in many areas. There are numerous geographical locations that are suitable for installation of large aerogenerators, which are to be found increasingly on the European landscape, particularly in Denmark and Germany. In the U.S., areas of California, Kansas, and West Texas are particularly well adapted to the installation of wind generating facilities. Some people regard the tall, graceful aerogenerators as ugly (others see them as graceful and picturesque), but they are certainly not as ugly as the landscape will become if massive climate warming occurs.

There is one big problem common to solar and wind energy—their intermittent nature. Solar energy works poorly when the sun does not shine and wind energy fails when the wind does not blow (although modern aerogenerators function at remarkably low wind speeds). Therefore, it is necessary to have reliable means of energy storage to provide for an even energy flow. Batteries for storage in major electrical systems would be too large and expensive. Possible alternatives include pumped water storage, kinetic energy in rapidly spinning flywheels, and production of hydrogen and oxygen by electrolysis of water with these two gases later recombined in fuel cells to produce electricity.

For some applications, most notably aircraft fuels, there are no realistic substitutes for carbon-based fuels. The challenge is to utilize such fuels with minimal addition of greenhouse gas carbon dioxide to the atmosphere. The best sustainable alternative for producing such fuels is to make them from biomass. It is possible to synthesize and utilize biomass fuels in a manner that is greenhouse-gas-neutral, that is, the carbon in the carbon dioxide released by their combustion came originally from the atmosphere by photosynthesis (as did the carbon in fossil fuels, but over a vastly longer time frame). Biomass is now used for liquid fuels in two major forms. One of these is ethanol made from the fermentation of sugar from grain starch or sugar from sugarcane. The other alternative is diesel fuel made by esterifying plant oils, particularly soybean oil. But these sources require a high-value raw material that is in demand for food and are economic only because of substantial government subsidies. Efforts to extract sugars for fermentation to alcohol from wood and crop byproduct sources including stalks, leaves, and straw have proven difficult and uneconomical.

The best alternative for preparing liquid fuels from biomass sources is thermochemical gasification, which produces a synthesis gas consisting of a mixture of carbon monoxide, CO, and elemental hydrogen, H_2. The proportion of H_2 can be increased by reacting CO with steam (H_2O). The CO and H_2 can be combined in various proportions to produce a wide range of fuels including methane, gasoline, jet fuel, and diesel fuel. The fraction of biomass that is consumed to generate the H_2 required to make synthetic fuels, a

process that generates one molecule of greenhouse gas CO_2 for each molecule of H_2 produced, can be greatly reduced by using hydrogen gas made by the electrolysis of water using renewable wind energy.

Other alternatives for energy production, conservation, and utilization are presented in Chapter 15. The most promising of these are summarized below:

- Use of water power without dams by means including turbines anchored in flowing river water and tidal energy sources.
- Generation of geothermal energy, especially from dry hot rock sources.
- Use of fossil fuels with underground carbon dioxide sequestration.
- Increased use of nuclear power with breeder reactors that produce more fissionable matter than they consume and destruction of long-lived fissionable radioactive transuranic elements such as plutonium by their use as reactor fuel.
- Conversion of rail systems from diesel-powered locomotives to electricity generated from renewable sources.
- Increased use of methane (synthetic natural gas) made from biomass sources as fuel for motor vehicles
- Storage of energy from intermittent renewable sources by means such as pumping hydrogen gas made by electrolysis of water into underground storage sites

---

This page titled 2.5.17.5: The Fourth Commandment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **17.5: The Fourth Commandment** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.17.6: The Fifth Commandment

CLIMATE CONDUCIVE TO LIFE ON EARTH MUST BE MAINTAINED AND ACCEPTABLE MEANS MUST BE FOUND TO DEAL WITH CLIMATE CHANGES THAT INEVITABLY OCCUR

The most plausible way that humans can ruin the global environment is by modifying the atmosphere such that global warming on a massive scale occurs. The most likely cause of such a greenhouse effect is release of carbon dioxide into the atmosphere from fossil fuel combustion as discussed in Section 10.6. Human activities are definitely increasing atmospheric carbon dioxide levels and there is credible scientific evidence that global warming is taking place. These phenomena and the climate changes that will result pose perhaps the greatest challenge for human existence, at least in a reasonably comfortable state, on the planet.

The Fifth Commandment is very much connected with the Fourth Commandment because so much of the increased atmospheric carbon dioxide levels are tied with energy and fossil fuel use. Other factors are involved as well. Destruction of forests (see the Sixth Commandment below)removes the carbon dioxide-fixing capacity of trees, and the decay of biomass residues from forests releases additional carbon dioxide to the atmosphere. Methane is also a greenhouse gas. Itis released to the atmosphere by flatulent emissions of ruminant animals (cows, sheep, moose), from the digestive tracts of termites attacking wood, and from anoxic bacteria growing in flooded rice paddies. Some synthetic gases, particularly virtually indestructible fluorocarbons, are potent greenhouse gases as well. The achievement of sustainability requires minimization of those practices that result in greenhouse gas emissions, particularly the burning of fossil fuels

Unfortunately, if predictions of greenhouse gas warming of Earth’s climate are accurate, some climate change inevitably will occur. Therefore, it will be necessary to adapt to warming and the climate variations that it will cause. Some of the measures that will have to be taken are listed below:

- Relocation of agricultural production from drought-plagued areas to those mademore hospitable to crops by global warming (in the Northern Hemisphereagricultural areas are likely to shift northward)
- Massive irrigation projects to compensate for drought
- Development of heat-resistant, drought-tolerant crops
- Relocation of populations from low-lying coastal areas flooded by rising sea levels caused by melted ice and expansion due to warming of ocean water
- Construction of sea walls and other structures to compensate for rising sea levels
- Water desalination plants using sea water and brackish groundwater to make up for reduced precipitation in some areas
- Increased utilization of saline water such as growing algae for biomass energy production in salt water instead of fresh water

This page titled 2.5.17.6: The Fifth Commandment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **17.6: The Fifth Commandment** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.17.7: The Sixth Commandment

The loss of Earth's biological productivity would certainly adversely affect sustainability and, in the worst case, could lead to massive starvation of human populations. A number of human activities have been tending to adversely affect biological productivity, but these effects have been largely masked by remarkable advances in agriculture such as by increased use of fertilizer, development of highly productive hybrid crops, and widespread irrigation. Some of the factors reducing productivity are the following:

- Loss of topsoil through destructive agricultural practices
- Urbanization of land and paving of large amounts of land area
- Desertification in which once productive land is degraded to desert
- Deforestation
- Air pollution that adversely affects plant growth

Biological productivity is far more than a matter of proper soil conditions. In order to preserve and enhance biological productivity, all five environmental spheres must be considered. Obviously, in the geosphere, topsoil must be preserved; once it is lost, the capacity of land to produce biomass is almost impossible to restore. Deforestation must be reversed and reforestation of areas no longer suitable for crop production promoted. (This is happening in parts of New England where rocky, hilly farmland is no longer economical to use for crop production.) In more arid regions where trees grow poorly, prairie lands should be preserved, desertification from overgrazing and other abuse prevented, and marginal crop lands restored to grass.

The hydrosphere may be managed in a way to enhance biological productivity. Measures such as terracing of land to minimize destructive rapid runoff of rainfall and to maximize water infiltration into groundwater aquifers may be taken. Watersheds, areas of land that collect rainwater and which may be areas of high biological productivity should be preserved and enhanced.

It is especially important that the atmosphere be maintained in a condition of climate conducive to high bioproductivity by minimization of global warming.

Management of the biosphere, itself, may enhance biological productivity. This has long been done with highly productive crops. The production of wood and wood pulp on forest lands can be increased—sometimes dramatically—with high-yielding trees, such as some hybrid poplars. Hybrid poplars from the same genus as cottonwoods or aspen trees grow faster than any other tree variety in northern temperate regions, so much so that for some applications they may be harvested annually. They have the additional advantage of spontaneous regrowth from stumps left from harvesting.

Proper management of the anthrosphere is essential to maintaining biological productivity. The practice of paving large areas of productive land should be checked. Factories in the anthrosphere can be used to produce fertilizers for increased biological productivity.

This page titled 2.5.17.7: The Sixth Commandment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

- **17.7: The Sixth Commandment** by Stanley E. Manahan is licensed CC BY-NC-SA 4.0.

2.5.17.8: The Seventh Commandment

MATERIAL DEMAND MUST BE DRASTICALLY REDUCED; MATERIALS MUST COME FROM RENEWABLE SOURCES, BE RECYCLABLE AND, IF DISCARDED TO THE ENVIRONMENT, BE DEGRADABLE

Reduced material demand is essential to sustainability. Fortunately, much is being done to reduce material demand and the potential exists for much greater reductions. Nowhere is this more obvious than in the communications and electronics industries. Old photos of rail lines from the early 1900s show them lined with poles holding 10 or 20 heavy copper wires, each for carrying telephone and telegraph communications. Now far more information can be carried by a single thread-sized strand of fiber optic material. The circuitry of a bulky 1948-vintage radio with its heavy transformers and glowing vacuum tubes has been replaced by circuit chips smaller than a fingernail. These are examples of dematerialization and also illustrate material substitution. For example, fiber optic cables are made from silica extracted from limitless supplies of sand whereas the conducting wires that they replace are made from scarce copper.

Wherever possible, materials should come from renewable sources. This favors wood, for example, over petroleum-based plastics for material. Wood and other biomass sources can be converted to plastics and other materials. From a materials sustainability viewpoint natural rubber is superior to petroleum-based synthetic rubber, and it is entirely possible that advances in genetic engineering will enable growth of rubber-producing plants in areas where natural rubber cannot now be produced.

Materials should be recyclable insofar as possible. Much of the recyclability of materials has to do with how they are used. For example, binding metal components strongly to plastics makes it relatively more difficult to recycle metals. Therefore, it is useful to design apparatus, such as automobiles or electronic devices, in a manner that facilitates recycling.

Some materials, by the nature of their uses, have to be discarded to the environment. An example of such a material is household detergent, which ends up in wastewater. Such materials should be readily degradable, usually by the action of microorganisms. Detergents provide an excellent example of a success story with respect to degradability. The household detergents that came into widespread use after World War II contained ABS surfactant (which makes the water “wetter”) that was poorly biodegradable such that sewage treatment plants and receiving waters were plagued with huge beds of foam. The ABS surfactant was replaced by LAS surfactant which is readily broken down by bacteria and the problem with undegradable surfactant in water was solved.

This page titled 2.5.17.8: The Seventh Commandment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.17.9: The Ninth Commandment

IT MUST BE ACKNOWLEDGED THAT THERE ARE RISKS IN TAKING NO RISKS

Some things for which there are no suitable substitutes are inherently dangerous. We must avoid becoming so risk adverse that we do not allow dangerous, but necessary activities (some would put sex in this category) to occur. A prime example is nuclear energy. The idea of using a“controlled atom bomb” to generate energy is a very serious one. But the alternative of continuing to burn large amounts of greenhouse-gas-generating fossil fuels, with the climate changes that almost certainly will result, or of severely curtailing energy use, with the poverty and other ill effects that would almost certainly ensue, indicates that the nuclear option is a good one.

So it is necessary to manage risk and to use risky technologies in a safe way. As discussed in Chapter 15, with proper design and operation, nuclear power plants can be operated safely and spent nuclear fuel can be processed safely. Modern technology and applications of computers can be powerful tools in reducing risks. Computerized design of devices and systems can enable designers to foresee risks and plan safer alternatives. Computerized control can enable safe operation of processes such as those in chemical manufacture. Redundancy can be built into computerized systems to compensate for failures that may occur. The attention of computers does not wander, they do not do drugs, become psychotic, or do malicious things (although people who use them are not so sure). Furthermore, as computerized robotics advance, it is increasingly possible for expendable robots to do dangerous things in dangerous areas where in the past humans would have been called upon to take risks.

Although the goal of risk avoidance in green chemistry and green technology as a whole is a laudable one, it should be kept in mind that without a willingness to take some risks, many useful things would never get done. Without risk-takers in the early days of aviation, we would not have the generally safe and reliable commercial aviation systems that exist today. Without the risks involved in testing experimental pharmaceuticals, many life-saving drugs would never make it to the market. Although risks must be taken judiciously, a total unwillingness to take risks will result in stagnation and a lack of progress in important areas required for sustainability.

2.5.17.9: The Ninth Commandment is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.5.17.10: The Tenth Commandment

EDUCATION IN SUSTAINABILITY IS ESSENTIAL; IT MUST EXTEND TO ALL AGES AND STRATA OF SOCIETY, IT MUST BE PROMULGATED THROUGH ALL MEDIA, AND IT IS THE RESPONSIBILITY OF ALL WHO HAVE EXPERTISE IN SUSTAINABILITY

Although the achievement of sustainability is the central challenge facing humanity, most people know pathetically little about it. The reader of this chapter belongs to a small fraction of the populace who have been exposed to the idea of sustainability. If asked, a distressingly large number of people would probably say that they know little about sustainability (and some would say that they do not even care or that they are even hostile to the concept!). Therefore, education is essential and a key to achieving sustainability.

Education in sustainability must begin early with children in primary school and should be integrated into curricula from kindergarten through graduate school. By providing containers for recyclables in grade schools, there is some small benefit from the waste paper, plastics, and aluminum cans collected, but a much greater benefit in the lessons of sustainability that those containers illustrate. Green chemistry should be part of the background of every student graduating with a university degree in chemistry and the principles of green engineering should be part of the knowledge base of every engineering graduate. But of equal —often greater—importance is the education of people in nontechnical areas in the principles of sustainability. Lawyers, political scientists, economists, and medical professionals should all graduate with education in sustainability.

A particular challenge is that of informing the general public of the principles of sustainability and of its importance. The general public has more choice in its sources of information than does the captive audience of a student body, so the challenge of informing them about sustainability is greater. In this respect the media and the internet have key roles to play. Unfortunately, relative to the large amounts of media time devoted to the salacious antics of some attention-seeking fools—matters that have virtually no relevance to the lives of everyday citizens—almost no air time is devoted to sustainability, which is highly relevant to the lives of all. Therefore, those who have an interest in, and knowledge of sustainability have an obligation to get the message out through the media and the internet.

2.5.17.10: The Tenth Commandment is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.5.17.11: Some Sensible Measures for Sustainability

Some ideas for measures to enhance sustainability are discussed here. Some of these are simple and readily implemented. Others are grandiose and even “far out.” Most have probably been suggested in one form or another by others or have been implemented in some places. Few require the invention of anything new. They are presented here with the idea of stimulating thought and discussion. Probably the reader can suggest other ideas for sustainability. Although most of these ideas are presented from the viewpoint of their potential implementation in the United States, most of them, or closely related measures, could be implemented in other countries as well.

The Methane Energy Economy

The basic idea of the methane energy economy is to greatly increase the role of natural gas as a fuel, particularly for transportation. Methane, CH_4, is the cleanest burning of the fossil fuels and produces the least greenhouse gas per unit energy generated. Significantly more energy can be stored in a pressurized tank of methane than is the case for elemental hydrogen. Largely regarded as a rapidly depleting source of fuel as recently as the turn of the century (2000), natural gas has emerged as an abundant energy source. There are two major reasons for this abundance. One is a revolutionary development of methods for gas extraction from hydraulically-fractured tight shale formations. These formations, previously inaccessible for natural gas extraction, are widely distributed in many countries including in the U. S. wide areas of Pennsylvania, Arkansas, Louisiana, and Texas (one of the richest deposits lies beneath the city of Fort Worth, Texas). A second factor is the construction of large depots for the export of liquified natural gas from countries that formerly burned it in flares as a waste byproduct of crude oil production. As of 2010it was projected that the U.S. had 100 years of natural gas reserves and might become an exporter of this fuel.

As depletable underground sources of methane become exhausted, methane can be made by gasification of coal and biomass and the byproduct carbon dioxide from solids gasification converted to methane by reacting CO_2 with hydrogen made from water electrolysis using electricity from renewable sources such as wind. Methane is readily distributed by pipeline, and a system for so doing is already in place in the U.S. and many other nations. Once properly installed underground, a pipeline is a less disruptive, highly reliable, means of transporting energy than is a high-voltage electrical line or transport of liquid fuels by train or barge. An important consideration is that the pipeline infrastructure must be very well designed, constructed, and maintained to avoid destructive blowouts. A very bad natural gas pipeline failure happened on September 9, 2010, with the rupture of a 30-inch natural gas pipeline in San Bruno, California. This devastating incident resulted in an explosion and massive fire that killed 8 people, destroyed 38 houses, damaged many more homes, and left a crater 12 meters deep, 51 m long, and 8 m wide.

Conversion to Highly Efficient Hybrid Vehicles

As noted above, methane is a very sustainable energy source that can replace most existing applications of fossil fuels. One of the most attractive options is to use methane to power locomotives, trucks, and other vehicles. The Clean Air Power Company has designed highly energy-efficient truck engines that get 90% of their energy from methane fed into the engine with intake air and that are ignited by a small amount of diesel fuel injected at the peak of the compression stroke.[5] This is an application of the stratified-charge ignition concept that enables internal combustion engines to run on a fuel/air mixture that is too lean (fuel-deficient) to ignite with a spark but can be ignited by a small fuel-rich ignition zone created by injecting fuel directly into a small region of the combustion chamber (in a spark-ignited engine directly onto the spark plug).

The ultimate in automobile fuel economy would be a plug-in hybrid vehicle with a natural gas fueled stratified-charge ignition engine (see above). The battery on such a vehicle could be charged by plugging into a source of electricity and with a full charge could be driven several tens of miles distance before the internal combustion engine would need to be engaged. When needed to charge the battery, the internal combustion engine could be run at optimum speeds for comparatively long periods of time for maximum efficiency and minimum pollution. One problem with hybrid vehicles is the need for the internal combustion engine to start immediately in cold weather to provide heat for the vehicle heater. This problem may be overcome with a phase-change material held as a heated liquid in an insulated container and that releases latent heat as it solidifies.[6]

Diverting the Mississippi

In Chapter 8 a scheme was presented to utilize the vast resource of the Mississippi River by diverting part of its flow from the Gulf of Mexico to water-deficient regions of the southwesternU.S. and northern Mexico. The scheme would impound a fraction of the

river discharge into a contained wetlands area at the mouth of the river, allow natural processes to purify the water, and use wind energy to pump the purified water as much as 2000 miles as far as southern California for municipal water supply and irrigation. Major sustainability aspects of this scheme are the following:

• Plant growth in the constructed wetlands impoundment would remove nutrients from the water thereby reducing their discharge into the Gulf of Mexico where they cause a eutrophied “dead zone.”

• Plant biomass harvested from the constructed wetlands could be used for synthetic fuels production.

• River sediment collected in the impoundment would aid in coastal restoration.

• Water transported to arid regions would reduce current excessive demand for water from the Colorado and Rio Grande rivers.

• Fish and freshwater shrimp grown in the waterway and impoundments along its course could provide a significant source of protein

Three Day Per Week Mail Delivery

In the modern electronic age, communication by conventional mail is relatively less important than it was in the pre-computer age. Significant cost and energy savings could be achieved by reducing mail delivery to three days per week, Monday, Wednesday, Friday on some routes and Tuesday, Thursday, Saturday on alternate routes. Such a system would need to be implemented gradually to avoid layoffs of personnel.

Rail Connections to and within All Major Airports

Ground transportation systems to many major airports are cumbersome, slow, and wasteful oftime and energy. Each large airport should have an integrated rail system linking terminals to the rail system serving the city and, within the airport area itself, parking areas, car rental locations,hotels and other facilities. Such a totally integrated system would save energy and time and significantly reduce pollution in the airport complex.

Fertilization of Algae Beds with Exhaust Carbon Dioxide

This proposal is to use combustion exhaust gas from powerplant furnaces to enrich algae cultures with carbon dioxide. The efficiency of algal media to sequester carbon dioxide is enhanced by the ability of algae to make media basic in the production of biomass ($\{CH_2O\}$) by the following photosynthesis reaction

$$HCO_3^- + H_2O + h\nu \rightarrow (CH_2O) + O_2 + OH^-$$

Such a system would further purify the exhaust gas by sequestering acid gas, particularly SO_2, and particulate matter

Nuclear Fuel Reprocessing

Burying spent nuclear fuel in a permanent disposal repository is a wasteful practice inconsistent with the principles of sustainability and green chemistry. A much better practice is to reprocess the spent fuel as explained in the discussion of nuclear energy in Section 15.10. The relatively short-lived fission products can be isolated for burial and will decay to harmless levels within a few centuries. The longer-lived transuranic elements, several of which are useful fissionable fuels can be treated by neutron bombardment, a process called transmutation. As an interim solution, relatively long term storage of spent fuel elements at reactor sites (up to 100 years, for example) enables decay of a very large fraction of the initial radioactivity making it easier to reprocess the fuels. Such long term storage also allows time for the development of improved recycling technologies and breeder reactors that can use spent fuels.

Extraction of Heat from Wastewater

Wastewater is a significant source of heat that can be extracted with a heat pump for heating homes and commercial buildings. The concept involves collecting filtered sewage in a storage tank and pumping heat from the water with a heat pump. Subsequently, the water, cooled by several degrees, can be discharged to the sewer. The same system can be used for cooling in which case heat can be pumped into the water before it is discharged

2.5.17.11: Some Sensible Measures for Sustainability is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.5.17.12: The Eighth Commandment

THE PRODUCTION AND USE OF TOXIC, DANGEROUS, PERSISTENT SUBSTANCES SHOULD BE MINIMIZED AND SUCH SUBSTANCES SHOULD NOT BE RELEASED TO THE ENVIRONMENT; ANY WASTES DISPOSED TO DISPOSAL SITES SHOULD BE CONVERTED TO NONHAZARDOUS FORM

The most fundamental tenet of green chemistry is to avoid the production and use of toxic, dangerous, persistent substances and to prevent their release to the environment. With the caveat that it is not always possible to totally avoid such substances (see the Ninth Commandment below)significant progress has been made in this aspect of green chemistry. Much research is ongoing in the field of chemical synthesis to minimize toxic and dangerous substances. In cases where such substances must be used because no substitutes are available, it is often possible to make minimum amounts of the materials on demand so that large stocks of dangerous materials need not be maintained.

Many of the environmental problems of recent decades have been the result of improperly disposed hazardous wastes. Current practice calls for placing hazardous waste materials in secure chemical landfills. There are two problems with this approach. One is that, without inordinate expenditures, landfills are not truly “secure” and the second is that, unlike radioactive materials that do eventually decay to nonradioactive substances, some refractory chemical wastes never truly degrade to nonhazardous substances. Part of the solution is to install monitoring facilities around hazardous waste disposal facilities and watch for leakage and emissions. But problems may show up hundreds of years later, not a good legacy to leave to future generations.

Therefore, any wastes that are disposed should first be converted to nonhazardous forms. This means destruction of organics and conversion of any hazardous elements to forms that will not leach into water or evaporate. A good approach toward this goal is to co-fire hazardous wastes with fuel in cement kilns; the organics are destroyed and the alkaline cement sequesters acid gas emissions and heavy metals. Ideally, hazardous elements, such as lead, can be reclaimed and recycled for useful purposes. Conversion of hazardous wastes to nonhazardous forms may require expenditure of large amounts of energy (see the fourth commandment, above)

This page titled 2.5.17.12: The Eighth Commandment is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Stanley E. Manahan.

2.5.17.13: Literature Cited

1. Ehrlich, Paul R., *The Population Bomb*, Ballantine Books, New York, 1968

2. Simon, Julian, *Hoodwinking the Nation*, Transaction Publishers, Somerset, NJ, 1999.

3. Diamond, Jared, *Collapse: How Societies Choose to Fail or Succeed*, Viking, New York, 2005.

4. Kolstad, Charles D., *Environmental Economics*, Oxford University Press, Oxford, UK, 2010.

5. Davy, M., R. L. Evans, and A. Mezo, "The Ultra Lean Burn Partially Stratified Charge Natural Gas Engine," presented at the 9th International Conference on Engines, and Vehicles, Naples, Italy, September 2009, seehttp://papers.sae.org/2009-24-0115.

6. Kenisarin, Murat and Khamid Mahkamov, "Solar Energy Storage Using Phase Change Materials," *Renewable and Sustainable Energy Reviews*,11, 1913-1965 (2007).

2.5.17.13: Literature Cited is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.5.17.14: Questions and Problems

Access to and use of the internet is assumed in answering all questions including general information, statistics, constants, and mathematical formulas required to solve problems. These questions are designed to promote inquiry and thought rather than just finding material in the text. So in some cases there may be several "right" answers. Therefore, if your answer reflects intellectual effort and a search for information from available sources, your answer can be considered to be "right.

1. The U.S. Geological Survey posts prices in metals corrected for inflation from 1959 through 1998 athttp://minerals.usgs.gov/minerals/.../metal_prices/. Other internet sources give current prices of metals. Using these resources, look up past and current prices of copper, chromium, nickel, tin and tungsten, the metals in the historic Ehrlich/Simon wager to see how the wager would have turned out in recent years. Explain your observations and suggest future trends.

2. At the beginning of the chapter were stated six "great challenges to sustainability." Using adiagram, if appropriate, suggest how these challenges are very much interrelated.

3. Much of Earth's population is dependent upon seafood for its protein. Suggest how provisionof adequate food by this route might be affected by the challenges of adequate energy supply as related to contamination of Earth's environment with toxic and persistent substances.

4. Suggest scenarios by which the people of your country might be adversely affected and its population might even decline because of declines in key environmental support systems. Suggest how these problems might be avoided in the future by actions taken now.

5. Look up telecommuting on the internet and explain how it might fit with the ten commandments of sustainability.

6. In a quote attributed to the British sociologist Martin Albrow, sociospheres consist of "distinct patterns of social activities belonging to networks of social relations of very different intensity, spanning widely different territorial extents, from a few to many thousands of miles." Suggest how sociospheres relate to the anthrosphere. What is your sociosphere? How might sociospheres be important in sustainability.

7. It has been estimated that U.S. coal resources can provide for up to three centuries of U.S.energy needs and could provide energy for much of the world. Is this a sustainable source ofenergy? Explain.

8. Explain how unsustainability may be the result of depletion of resources, environmentalpollution, or a combination. Give an example in which environmental pollution reducesresource availability.

9. From news events of the last five years, cite evidence that global climate change is in factoccurring. Cite evidence to the contrary.10. Currently, population growth tends to occur in coastal areas. Suggest how global warmingmight reverse that trend.11. Prior to European settlement, vast areas of the U.S. Great Plains supported huge herds of bisonthat provided the base for a viable Native American population. Given the erratic climate ofthat region, depleted groundwater for irrigation, and the aversion of many people to dwell onthe "lone prairie," some authorities have suggested that these areas revert to a bison-basedsystem, sometimes called the "buffalo commons." Suggest how such a system might be viableand sustainable and give arguments against it.12. Is the term "secure chemical landfill" an oxymoron? What are the alternatives to suchlandfills?13. Among the ideas for sustainability discussed in Section 17.12, might be added a suggestion forgreatly increasing the numbers of telecommuters. What are telecommuters? How mightincreasing their numbers enhance sustainability?14. Some countries in Asia have constructed new airports with exemplary rail connections. Citeand discuss several examples.

2.5.17.14: Questions and Problems is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.5.17.15: Supplementary References

Assadourian, Erik and Muhammed Yunus, *State of the World 2010: Transforming Cultures from Consumerism to Sustainability*, W.W. Norton and Company, New York 2010.

Ayres, Robert U., and Benjamin Warr, *The Economic Growth Engine: How Energy and Work Drive Material Prosperity*, Edward Elgar, Northampton, MA, 2009.

Brand, Stewart, *Whole Earth Discipline: Why Dense Cities, Nuclear Power, Transgenic Crops, Restored Wildlands, and Geoengineering are Necessary*, Penguin, New York, 2010.

Dresner, Simon, *The Principles of Sustainability*, 2nd ed., Earthscan Publications, London, 2008.

Hawken, Paul, Amory Lovins, and L. Hunter Lovins, *Natural Capitalism: Creating the Next Industrial Revolution*, Back Bay Books, Boston, 2008.

Lankey, Rebecca L., and Paul T. Anastas, *Advancing Sustainability through Green Chemistry and Engineering*, American Chemical Society, Washington, DC, 2002.

Owen, David, *Green Metropolis: Why Living Smaller, Living Closer, and Driving Less are the Keys to Sustainability*, Riverhead Hardcover, New York, 2009.

Vallero, Daniel A., *Sustainable Design: The Science of Sustainability and Green Engineering*, Wiley, Hoboken, NJ, 2008.

2.5.17.15: Supplementary References is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

Back Matter

Detailed Licensing

Detailed Licensing

2.6: Toxicology MSDT

One of the challenges of recruiting undergraduates into the toxicology graduate research programs is a lack toxicology training programs and courses at the undergraduate level. Course content developed through this program are accessible to the general public. The intent is to provide content where other undergraduate students, K-12 educators, and the interested public encounter content providing an introduction to the discipline of toxicology.

This page titled 2.6: Toxicology MSDT is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

Front Matter

Licensing
Preface

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

Preface

What is Toxicology?

Toxicology is the science of characterizing the effects of poisons (toxicants) on living organisms. It is a broad interdisciplinary scientific field which includes life sciences, chemistry, and engineering among others. Toxicologists work towards the protection of human and animal health, and of the environment.

Why ToxMSDT?

The Toxicology Mentoring and Skills Development Training Program is a 2 yr pilot program funded by the National Institutes of Health. The goal of the program is to support educational activities that complement and/or enhance the training of a diverse workforce to meet the nation's biomedical, behavioral and clinical research needs. Toxicology is an essential component of the nation's biomedical research enterprise. However, there is a critical lack of underrepresented populations in toxicology.

Iowa State University, Tuskegee University, and the Ohio State University conceived this unique Toxicology Mentoring and Skills Development Training program targeting undergraduate underrepresented populations in order to build a pipeline entering graduate school and the toxicology workforce. Core concepts of the program are structured mentoring activities, skills development, and public outreach in toxicology.

Nuts and Bolts of ToxMSDT

This is a one-year long toxicology mentoring and skills development training program. For the structured mentoring component, mentees are matched 1:1 with mentors primarily from industry, government, and NGOs. Mentors are trained to provide a holistic mentoring experience to mentees. Mentees shadow mentors at their places of work in addition to other face-to-face activities at Iowa State University, the annual Society of Toxicology meeting, and at Tuskegee University.

The Skills Development component consists of didactic training at the inaugural workshop held at Iowa State University, completion of six short e-modules covering core competencies of toxicologists, and knowledge gained through shadowing activities, and presentations by mentees to enhance their communication skills.

One of the challenges of recruiting undergraduates into the toxicology graduate research programs is a lack toxicology training programs and courses at the undergraduate level. Course content developed through this program are accessible to the general public. The intent is to provide content where other undergraduate students, K-12 educators, and the interested public encounter content providing an introduction to the discipline of toxicology.

- **Preface** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.1: Pathophysiology

Pathophysiology is the study of the physical and functional changes that occur during a disease process. In this e-module you will learn about the concept of pathophysiology, types of toxicity, repair and adaptation, and patterns of toxic injury.

2.6.1.1: What is Pathophysiology?

2.6.1.2: Outcomes of Targeted and Non-Targeted Toxicity

2.6.1.3: Cellular Response to Toxicant-Induced Injury

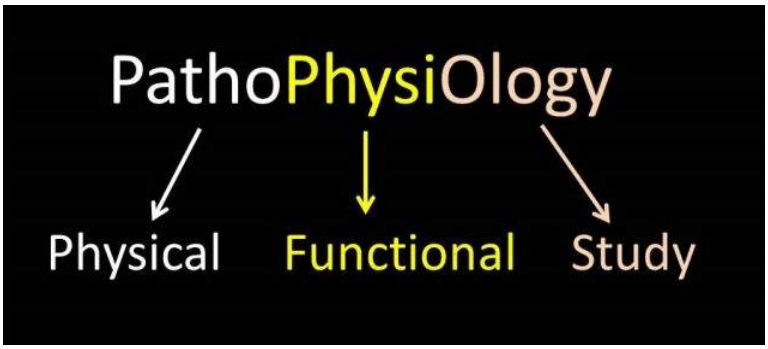
2.6.1.4: Repair and Adaptation

2.6.1.5: Patterns of Toxic Injury

2.6.1.6: Final Evaluation

This page titled 2.6.1: Pathophysiology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.1.1: What is Pathophysiology?



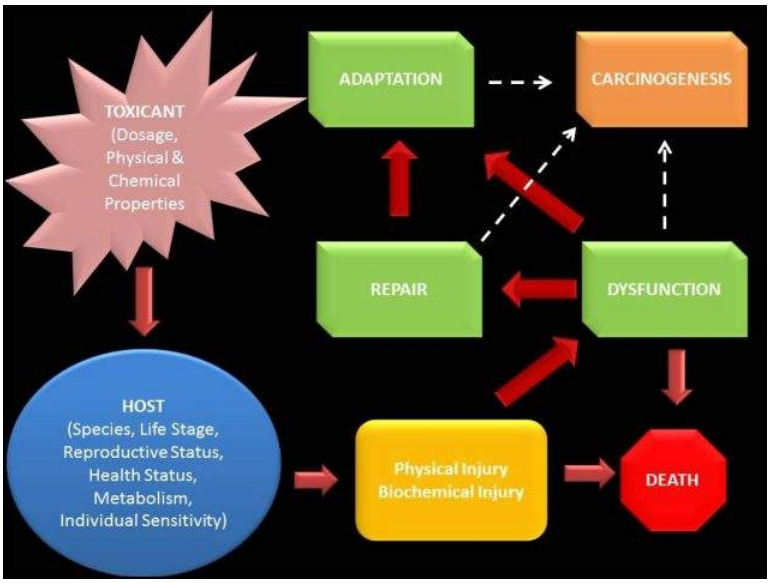
Pathophysiology describes the changes that occur during a disease process, with “patho-“ referring to the physical changes that are observed and “physio-“ referring to the functional processes or mechanisms that occur during a disease process. In toxicology, pathophysiology encompasses the biochemical and physical alterations that occur upon exposure of an individual (generally termed the “host”) to harmful amounts of a toxicant.







In toxicology, pathophysiology takes into account how the characteristics of the toxicant (for example, dosage, physical properties and chemical properties) and the characteristics of the host (including species, life stage, health/reproductive status, metabolism, and individual sensitivity) interact to produce physical and/or biochemical changes in the host. Pathophysiology also encompasses the host response to the effects of a toxicant. With acutely lethal intoxications, the physical and chemical injury may be sufficient to cause rapid death of the organism. In non-lethal toxic exposures, toxicant-induced injury results in dysfunction of cells, tissues and/or organs that may persist or that may progress to death. Persistent toxic injury that does not result in death generally leads to attempts at repair of toxicant-induced damage. With some toxicants, the host is able to develop strategies to adapt to continued exposure to toxicants. Dysfunction, repair and adaptive processes that occur in response to exposure to certain toxicants may trigger development of unregulated cell growth leading to tumor formation in a process termed carcinogenesis.



In many cases the toxicant is the unchanged xenobiotic to which the host was exposed, but in some cases the xenobiotic itself may be relatively innocuous and requires bioactivation to more toxic metabolites before toxic effects occur. A variety of endogenous systems have evolved to mitigate the effects of many toxicants and/or their metabolites. However, when these systems fail or when the dose of toxicant exceeds the capacity of the system to neutralize the toxic effects, poisoning occurs. The clinical syndrome associated with a poisoning is referred to as toxicosis.

Most toxicants exert their effects on specific molecules, tissues or organs based on the physical and chemical makeup of the toxicant as well as the absorption, distribution and metabolism of the toxicant within the body.

Toxicants such as some strong acids or alkalis (e.g. concentrated hydrochloric acid) are not systemically absorbed, so are limited to causing local injury upon contact with skin, eyes or mucous membranes. Toxicants that are ingested and absorbed from the gastrointestinal tract are shuttled via the portal vein to the liver, where they may cause direct injury, where they may be bioactivated to toxic metabolites, or where they may be detoxified before they reach the general circulation (a process termed “first pass effect”). Inhaled toxicants such as smoke may cause local tissue injury due to irritants or corrosive components as well as systemic intoxication from toxic gases such as carbon monoxide or cyanide.

Topic 1: Key Points
In this section, we explored the following main points:

- 1: Pathophysiology is the study of the physical and functional changes that occur during a disease process.

- 2: Toxic insults can result in physical and biochemical alterations that may lead to cellular dysfunction, repair, adaptation, carcinogenesis and/or death.

This page titled 2.6.1.1: What is Pathophysiology? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **1.1: What is Pathophysiology?** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.1.2: Outcomes of Targeted and Non-Targeted Toxicity

Learning Objectives

After completing this lesson, you will be able to:

- 1: Discuss the various effects that toxicants have on target molecules and how these effects result in injury to the host.
- 2: Discuss the pathogenesis of the rash that occurs following exposure to poison ivy.

The interaction of toxicants with host molecules may lead to the dysfunction or destruction of the target molecule, or it may result in formation of adducts that the immune system identifies as “foreign,” triggering immune responses against these neoantigens.

Target Molecule Dysfunction

Target molecule dysfunction is a common mechanism by which xenobiotics, particularly drugs, exert their effects; remember it is the dose of xenobiotic that determines whether the effect will be therapeutic (pharmacologic) or harmful (toxicologic). Target molecule dysfunction may occur through activation of cellular membrane receptors, resulting in over-stimulation of some cellular function. For instance, the toxic effects of methomyl, a carbamate insecticide, include over-stimulation of cells of excretory glands, resulting in excessive salivation, excessive tear formation and excessive secretion of mucus by goblet cells within the respiratory tract.

Conversely, other toxicants may inhibit or impede the action of cellular receptors; the resulting clinical effects will depend on the type of receptor affected and what action is impeded. For example, there are channels in nerve cell membranes that allow sodium to pass into and out of the cell; when exposed to pyrethrins, insecticides extracted from chrysanthemum flowers, these channels are unable to close, which results in excessive stimulation seen as muscle twitches, tremors and convulsions. However, when these channels are exposed to tetrodotoxin, the infamous puffer fish toxin, these channels are unable to open, which prevents stimulation of muscle and results in paralysis.



Figure 2.6.1.2.1: Pufferfish (fugu) is a delicacy in some countries, but if not prepared correctly it can cause paralysis and even death.

Toxicants may induce target molecule dysfunction by altering protein structure such that the protein is no longer functional, resulting in disruption of membrane protein channels, interference with transmembrane signaling or loss of enzyme function. Many of these types of effects involve the toxicant or its metabolite binding to reactive moieties on the protein molecule; the sulfhydryl or thiol (S-H) moiety is particularly susceptible to binding with other reactive compounds. Toxicant-induced alteration of DNA structure can lead to mispairing of nucleotides during mitosis, with potential effects ranging from altered protein synthesis to initiation of carcinogenesis.



Figure 2.6.1.2.2: DNA strand breaks may result in cell death, genetic mutation or development of cancers

Neoantigen formation results when a xenobiotic or its metabolite **binds** to a larger protein to form a novel molecule that elicits an immune response. Molecules that trigger an immune response upon binding to carrier proteins are termed **haptens,** and the process of neoantigen formation in this manner is termed **haptenization**. Neoantigens can trigger humoral immune responses resulting in the development of antibodies that can trigger acute **allergic** reactions such as hives or anaphylaxis. Neoantigens that trigger cellular-mediated immune responses cause injury to specific tissues or organs such as skin, liver or blood vessels in a process termed **autoimmunity.**

DID YOU KNOW?

The rash caused by poison ivy (Toxicodendron spp.) is caused by urushiol, an oily mixture of chemicals called catechols, in the sap of the plant. Upon exposure to human skin, urushiols bind to membranes on skin cells and serve as haptens, changing the shape of proteins in the membranes. The body's immune cells no longer recognize these skin cells as normal parts of the body and mount an immune response, resulting in inflammation, itching, blisters, swelling and redness at the site of urushiol contact. In addition to local irritation, serious, systemic reactions to urushiol can occur if the leaves are ingested, or if smoke from burning poison ivy is inhaled.



Figure 2.6.1.2.3: Skincontact with leaves of poison ivy can result in a blistering rash.

Topic 3: Key Points

In this section, we explored the following main points:

- 1: Toxicant-induced target molecule dysfunction can occur through activation or inhibition of cellular receptors, denaturing of membrane proteins, and destruction of target molecules.

- 2: Haptenization results in the formation of neoantigens that can trigger immune responses against cells and tissues of the body, resulting in allergic or autoimmune reactions.

    - The rash caused by poison ivy is an autoimmune reaction against skin cells whose membranes have bound to the toxicant urushiol from the plant.

? Knowledge Check

1. On a protein molecule highly reactive toxicant molecules have a predilection for_____.

    sulfhydryl (S-H) moiety

    Haptenization

    dose

Answer

sulfhydryl (S-H) moiety

2. The therapeutic or toxic effect of a xenobiotic is entirely dependent on its_____.

    sulfhydryl (S-H) moiety

    Haptenization

dose

Answer

dose

3. A xennobiotic binding a larger protein molecule, resulting in an immune response.

sulfhydryl (S-H) moiety

Haptenization

dose

Answer

Haptenization

This page titled 2.6.1.2: Outcomes of Targeted and Non-Targeted Toxicity is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

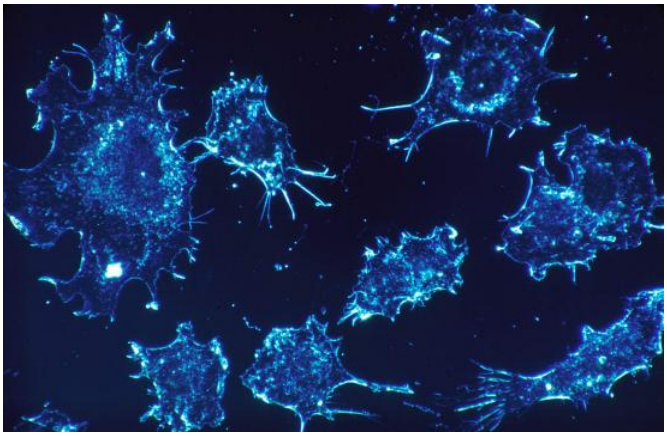
- **1.3: Outcomes of Targeted and Non-Targeted Toxicity** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.1.3: Cellular Response to Toxicant-Induced Injury

Learning Objectives

After completing this lesson, you will be able to:

- 1: Describe the range of cellular injury that may occur following exposure to a toxicant.
- 2: Discuss the major mechanisms of toxicant-induced cell death.
- 3: Compare the processes of necrosis and apoptosis in terms of inciting causes and the cellular changes that occur with each.



Toxicants can exert a variety of effects at the molecular level that have significant repercussions at the cellular, tissue and organ levels. The effects of a toxic exposure can range from reversible cellular dysfunction to irreversible cellular injury to cell death, all of which can alter normal organ function and have significant impact on the health and well-being of the body as a whole. The ability of cells, tissues and organs to overcome the effect of a toxicant through repair and/or adaptation will dictate the ultimate outcome of a toxic exposure.



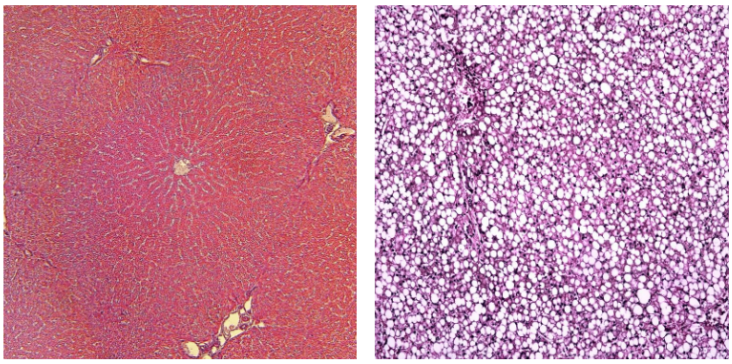


Figure 2.6.1.3.1: Lipid accumulation in the liver comparing normal liver (left) to liver with steatosis (center). In the photomicrographs at the right, the white spaces are areas of lipid accumulation.

Following toxic insult, cells have a limited repertoire of responses. Nonlethal cell injury may lead to cellular degeneration, seen microscopically as swelling of the cells. Toxicant-injured cells may accumulate water, lipid, pigments, glycogen or metabolic waste products due to impairment of normal maintenance functions. Accumulation of lipids in hepatocytes, termed **steatosis** or **hepatic lipidosis**, is a common toxic effect seen in cases of alcohol-related liver disease. Degenerative changes in cells are often reversible if the inciting cause is removed. When cell injury proceeds beyond the self-repair capability of the cell, cell death ensues.

Mechanisms of Cell Death

Major mechanisms of toxicant-induced cell death include disruption of cell membrane structure and/or function, loss of cellular maintenance functions, and impairment of cellular energy production. Loss of membrane structural or functional integrity can result in uncontrolled passage of water, ions and other compounds into or out of the cell. The subsequent loss of normal cytosolic environment interferes with normal biochemical processes necessary for cell function and/or survival. Loss of ability to synthesize proteins and other macromolecules impedes maintenance of organelles and enzymatic pathways vital to cellular survival. Impairment of cellular energy production generally occurs when toxic effects alter mitochondrial function and/or structure and can lead to cell death due to failure to produce sufficient ATP to power essential cellular functions.

Necrosis

Necrosis is the term used to describe cell death due to irreversible injury. Necrotic cells undergo degenerative processes including swelling of organelles, loss of organelle function, oxidative and hydrolytic degradation of intracellular membranes and macromolecules by electophiles and free radicals, and, ultimately, **lysis** (loss of cellular constituents to surrounding tissues due to cell membrane rupture). Necrosis generally results in the generation of an inflammatory response as cellular components and free radicals that are released to the extracellular matrix attract inflammatory cells.

Apoptosis

In contrast, **apoptosis** (sometimes nicknamed “cell suicide” or “programmed cell death”) is a more orderly form of cell death. Apoptosis is an active process involving activation of specific enzymes which triggers the systematic fragmentation of cell constituents into blebs of cell membrane that pinch off of the main cell to form **apoptotic bodies**. During this fragmentation, the cell continues to produce energy and proteins, unlike necrosis where organelle and energy production cease prior to cellular fragmentation. The end result of apoptosis is numerous apoptotic bodies, each composed of a cellular membrane surrounding intact and functional cellular components. Apoptosis can be triggered by various forms of oxidative stress, particularly the presence of excessive oxygen-derived free radicals, due to excessive free radical generation and/or to lack or exhaustion of endogenous antioxidants. Because intracellular components are not spilled into the extracellular matrix, apoptosis generally does not incite an inflammatory response; instead the apoptotic bodies are removed by local phagocytes.

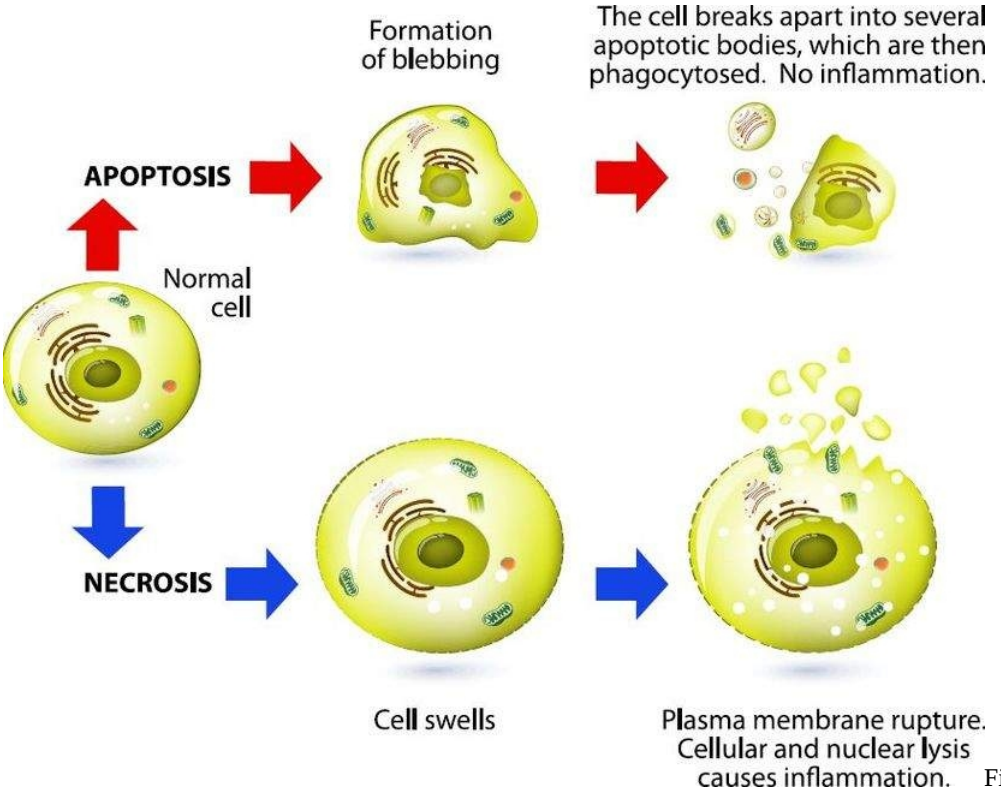


Figure 2.6.1.3.2: Apoptosis and Necrosis are two very different forms of cell death.

Topic 4: Key Points

In this section, we explored the following main points:

- 1: Toxicant-induced cellular injury can range from reversible cellular dysfunction to irreversible cellular injury to cell death.
- 2: Cellular responses to toxic injury may include cellular degeneration and accumulation of substrates within the cell.
    - Steatosis is lipid accumulation within hepatocytes and is a common toxic effect secondary to alcohol exposure.
- 3: Major mechanisms of cell death include disruption of cell membrane structure and/or function, loss of cellular maintenance functions, and impairment of cellular energy production.
- 4: Necrosis is cell death resulting from cessation of organelle function, degradation of intracellular structures and culminating in lysis of the cell and attraction of inflammatory cells.
- 5: Apoptosis is an active form of cell death whereby cellular function is maintained as the cell components are compartmentalized and packaged into apoptotic bodies that pinch off of the main cell.
    - Apoptosis can be triggered by oxidative stresses caused by oxygen-derived free radicals.
    - Apoptosis is not generally associated with an inflammatory response.

Knowledge Check

1. The accumulation of excess lipid within liver cells:

Answer

Steatosis

2. Cellular degeneration secondary to a toxic insult is seen microscopically as:

Answer

Cellular Swelling

3. Loss of organelle function, hydrolytic degradation of intracellular membranes and lysis of cells are characteristics of_____.

Answer

necrosis

4. The orderly decommissioning of cellular organelles without loss of energy and protein synthesis leading to fragmentation into membrane-bound packets is characteristic of_____.

Answer

apoptosis

This page titled 2.6.1.3: Cellular Response to Toxicant-Induced Injury is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **1.4: Cellular Response to Toxicant-Induced Injury** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

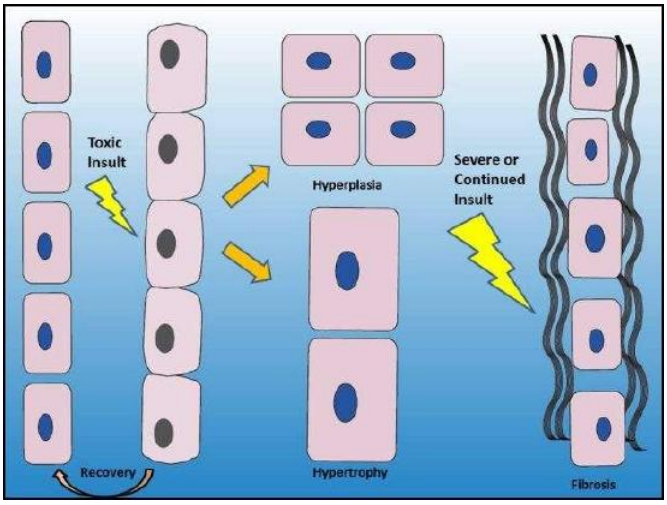
2.6.1.4: Repair and Adaptation

Learning Objectives

After completing this lesson, you will be able to:

1. Explain cell and tissue repair mechanisms and how toxicants may alter these processes.
2. Discuss adaptation mechanisms that may occur with exposures to toxicants.

Repair



The ability of cells to repair toxicant-induced damage plays a large role in the outcome of exposure to toxicants. Repair mechanisms include molecular repair such as reversal of thiol oxidation, removal of damaged units with replacement by newly synthesized units, and degradation followed by resynthesis of damaged structures. Cellular repair mechanisms include autophagy of damaged organelles and resynthesis/regeneration of damaged structures. Tissue repair mechanisms include active removal of damaged cells via apoptosis, regeneration of cells by **hyperplasia** (increase in cell numbers) or **hypertrophy** (increase in cell size), and resynthesis of extracellular matrix. When cellular or tissue repair is unable to restore original tissue architecture, production of extracellular connective tissue results in **fibrosis**.

Some toxicants can interfere with the normal cellular repair mechanisms, resulting in early cellular senescence and death. Inability to repair damage to DNA can result in disruption of normal protein synthesis or initiation of carcinogenesis. Other toxicants trigger exaggerated repair mechanisms which themselves pose a hazard to the survival of the individual. For instance, the herbicide paraquat causes lung injury mediated by oxygen-derived free radicals; because the lung is a highly oxygenated organ, this damage can be quite extensive as the ready availability of oxygen provides plenty of fuel for the snowballing generation of free radicals such as superoxide anions. The immediate effect of paraquat on the lung is damage to alveolar walls, resulting in stimulation repair mechanisms that cause intense, progressive fibrosis of the lung over several weeks, which ultimately leads to the death of the patient from asphyxia. Intense fibrosis is also seen in alcoholic cirrhosis of the liver, with normal liver parenchyma being replaced by fibrous connective tissue that impedes normal liver function.

ADAPTATION



Adaptation mechanisms have evolved at the cellular, tissue and organ levels that allow the individual to survive in the face of exposure to levels of toxicants that might otherwise result in serious or lethal injury. Adaptation requires time to develop, so generally occurs upon multiple exposures to levels of toxicants that are too low to cause severe acute injury. Adaptation mechanisms include alteration of toxicant delivery to target sites, decreasing reactivity of the target site to the toxicant, increasing local repair mechanisms, and development of compensatory mechanisms to mitigate toxicant-induced injury.

Decreasing toxicant delivery to target site may include decreasing toxicant absorption, detoxification of the toxicant before it can reach the target site, or binding of the toxicant with a neutral molecule. The opioid fentanyl is a powerful narcotic when injected, but it is quickly metabolized in the liver if ingested,

greatly reducing the amount that enters the systemic circulation and reaches the central nervous system; this “first pass effect” occurs with many substances and is often the reason why some drugs must be administered via injection rather than the oral route. Similarly, chronic ingestion of ethyl alcohol can result in the **induction** (increased expression) of the enzyme alcohol dehydrogenase in the gastrointestinal tract and liver, resulting in more alcohol being metabolized before it can reach the systemic circulation.

Some toxic heavy metals that reach the bloodstream become bound to metal-binding proteins called **metallothioneins**; because only unbound metals are free to react, metallothioneins prevent the toxic metals from interacting with their target sites. Many sea-dwelling mammals such as killer whales (Orca spp.) bind organic mercury to selenium-containing metallothioneins which allows them to accumulate large amounts of organic mercury that would be lethal to terrestrial mammals.

Adaptation by decreased reactivity to the target site is classically illustrated by the **tolerance** that can develop in people addicted to opioids such as heroin. Constant stimulation of opioid receptors results in **downregulation** (decreased expression) of those receptors and increased amounts of opioids required to stimulate remaining receptors.

Adaptive mechanisms may include increased function or number of cells to compensate for the loss of function of similar cells due to toxic insult. A variety of toxicants can cause injury to kidney tubules, and as a result, other kidney tubules may undergo hypertrophy and/or hyperplasia in an attempt to provide adequate kidney function.

Adaptive and repair mechanisms frequently occur simultaneously, resulting in characteristic lesions in some organs or tissues, For example, toxic hepatocellular injury from chronic alcohol exposure may result in nodular areas of regenerating hepatocytes (adaptation) surrounded by fibrous connective tissue formed to replace lost hepatocytes (repair), resulting in the characteristic “bubbly” look of hepatic cirrhosis.

Topic 5: Key Points

In this section, we explored the following main points:

- 1: Tissue repair mechanisms include removal of damaged cells via apoptosis, regeneration of cells by hyperplasia (increase in cell numbers) or hypertrophy (increase in cell size), and regeneration of extracellular matrix.
    - Fibrosis may result when tissue repair is incomplete.
- 2: Adaptation mechanisms may allow the host to survive in the face of continuous exposure to toxicants and include:
    - alteration of toxicant delivery to target sites.
    - decreasing reactivity of target site to toxicant
    - increasing local repair processes
    - compensatory mechanisms to mitigate toxicant-induced injury

Knowledge Check

1. An example of decreasing the reactivity of a target site to a toxicant

Answer

Downregulation of toxicant receptors

2. A regenerative response that results in an increased number of cells being produced

Answer

Hyperplasia

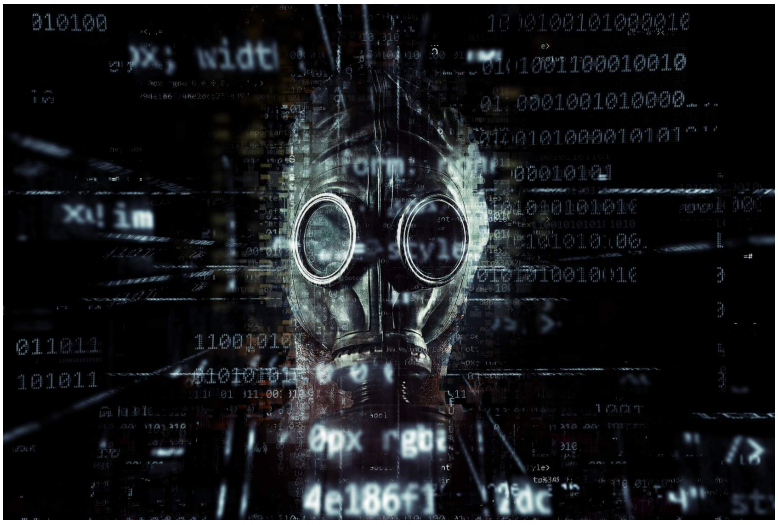
This page titled 2.6.1.4: Repair and Adaptation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **1.5: Repair & Adaptation** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.1.5: Patterns of Toxic Injury

Learning Objectives

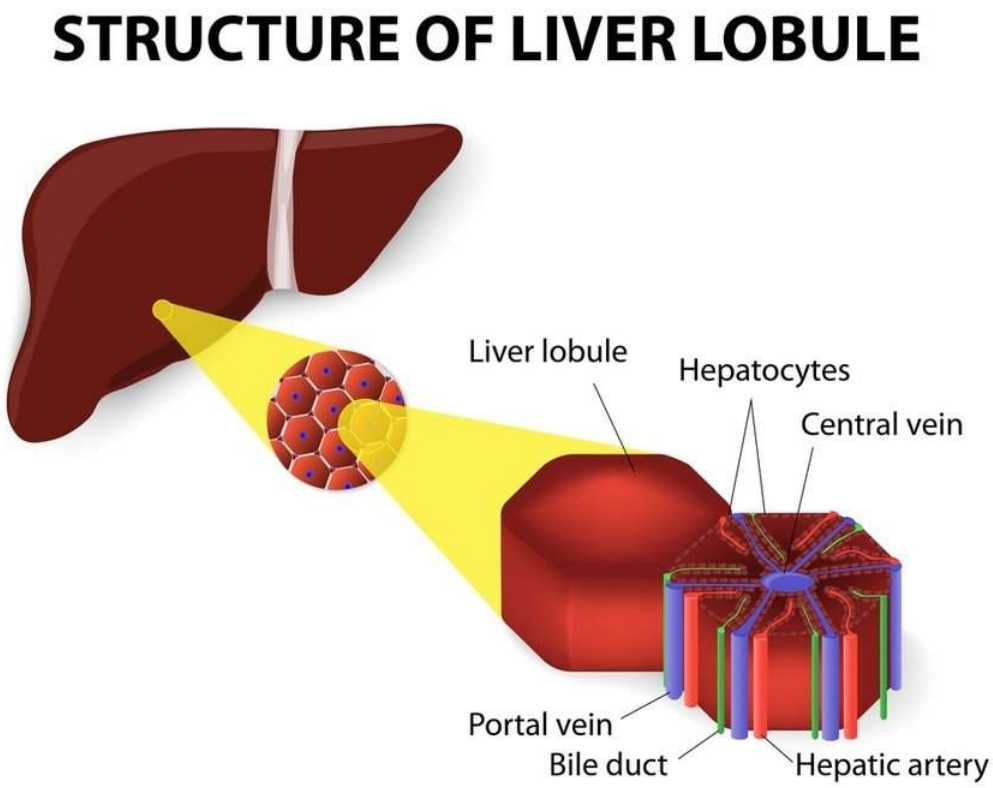
- 1: Explain how the structure and function of the liver relates to its susceptibility and responses to toxicants.
- 2: Discuss how the pattern of toxic injury seen in an organ may provide clues to the toxicant that is involved.



Toxicants can exert their effects systemically or by altering target tissues or organs, and sometimes the patterns of injury left behind can aid in the identification of an unknown toxicant, or at least allow the list of suspects to be narrowed down to a smaller number than the thousands of potential toxicants that exist. Because most absorbed toxicants are carried in the bloodstream, blood flow patterns often dictate which tissues may be affected. Within the tissues, variations in cell type and function can also influence the pattern of injury that can be seen with a given toxicant. Because they receive high volumes of blood flow, and because of their roles in metabolism and excretion of xenobiotics, the liver and kidney are particularly vulnerable to effect from systemically absorbed toxicants.

Patterns of Hipatic Injury

The microscopic appearance of the liver is that of a hexagonal lobule. Blood enters the parenchyma from the hepatic artery and portal vein situated, along with a bile duct, at an area termed the **portal triad** situated at each of the six vertices of the lobule (**periportal** regions). From the triad region, the blood passes through endothelial-lined **sinusoids** that separate anastomosing sheets of hepatocytes before exiting via the terminal hepatic, or **central** vein (**centrilobular** region). Physiologically, the liver has highly oxygenated blood entering at **zone 1**, which is roughly equivalent to the periportal areas, surrendering its oxygen as it filters through the sinusoids, and reaching **zone 3** (centrilobular region) as poorly oxygenated blood that exits via the central vein. **Zone 2**, also called the**midzonal** region, is the intermediate area between zones 1 and 3.



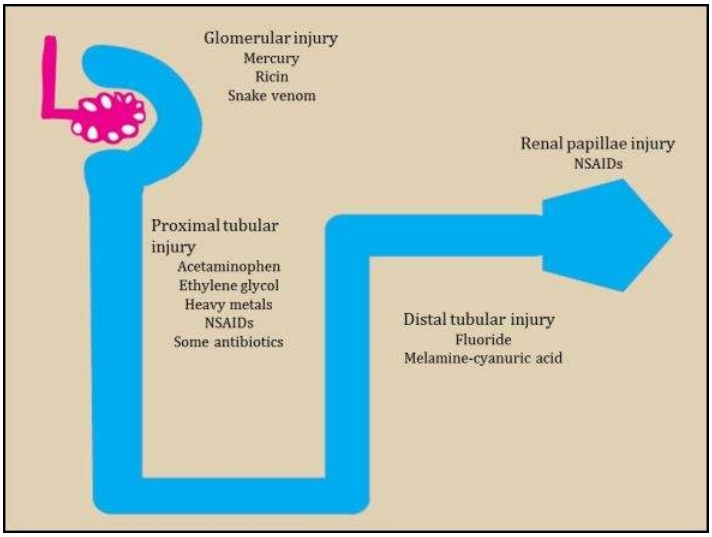
Within the different zones of the liver, hepatocytes vary in their physiologic functions, with those in zone 1 being more efficient in oxidative metabolism and zone 3 hepatocytes being efficient at xenobiotic biotransformation. Hepatocytes in zone 1 are the first to be exposed to toxicants that enter the liver; if those toxicants are directly injurious to hepatocytes (e.g. white phosphorus), the pattern of cell injury will be periportal. Toxicants requiring bioactivation to cause injury (e.g. acetaminophen) will generally cause zone 3 (centrilobular) hepatic injury, since this area contains higher levels of biotransforming enzymes. This area is also at risk for hypoxic injury due to toxicants that alter oxygen delivery to cells (e.g. carbon monoxide). Because of these vulnerabilities, centrilobular injury is the most common form of toxicant-induced hepatic injury. **Massive** liver necrosis affects entire hepatic lobules, and has been associated with exposure to a variety of toxicants including acetaminophen, aflatoxin, blue-green algae, and hepatotoxic mushrooms. **Piecemeal** necrosis is a less commonly seen form of liver injury wherein scattered individual hepatocyte necrosis or apoptosis occurs along the limiting plate between portal triads; this form of hepatic injury has been associated with immune-mediated processes such as is seen with non-steroidal anti-inflammatory drug-induced hepatopathy.

Zone 1 (Periportal)	Zone 2 (Midzonal)	Zone 3 (Centrilobular)
Iron Salts	Furans	Acetaminophen
White Phosphorus	Ngaione (*Myoporum* spp.)	Aflatoxin
	Beryllium	Microcystin (Blue-Green Algae)
		Ricin (Castor Bean)

Patterns of Renal Injury

The nephron is the functional unit of the kidney, consisting of the renal corpuscle (**Bowman's capsule** and **glomerulus**), **proximal tubule**, **loop of Henle**, and **distal tubules**. Toxicants that are directly injurious may cause damage to the glomerular structures or to the anterior portion of the proximal tubule. Toxicants that require bioactivation generally cause injury to the more distal section of

the proximal renal tubules, as this is where the majority of biotransforming enzymes occur. As the glomerular filtrate passes through the loop of Henle and distal tubules and becomes more concentrated, toxicants that were too dilute to affect earlier renal structures may cause injury in these more distal sections of the kidney. Toxicants that decrease renal blood flow (e.g. non-steroidal anti-inflammatory drugs) can also cause injury to more distal tubules, collecting ducts and renal papillae as these regions receive less blood flow than more proximal structures.



Topic 6: Key Points

In this section, we explored the following main points:

- 1: Toxic injury to a tissue or organ will depend largely on the pattern of blood flow to that tissue or organ.
    - Cells in areas that normally have low oxygenation are at increased risk of injury from toxicants that reduce blood flow or oxygen delivery.
- 2: Variations in cell type and function can affect the pattern of toxicant-induced injury.
    - Cells with large capacity for metabolism of xenobiotics will be at higher risk for injury by toxicants requiring bioactivation.
- 3: Identification of the pattern of toxicant-induced injury can provide clues to the types of toxicants that may have caused the injury.

Knowledge Check

1. Which of the following patterns of injury would be consistent with a toxicant that is directly injurious to liver and kidney cells?

    Centrilobular hepatic injury, distal tubular renal injury

    Centrilobular hepatic injury, proximal tubular renal injury

    Periportal hepatic injury, distal tubular renal injury

    Periportal hepatic injury, proximal tubular renal injury

Answer

Periportal hepatic injury, proximal tubular renal injury

2. Which features make the centrilobular region of the liver lobule more susceptible to toxic insult?

Relatively low levels of biotransformation enzymes and relatively low oxygenation

Relatively high levels of biotransformation enzymes and relatively low oxygenation

Relatively low levels of biotransformation enzymes and relatively high oxygenation

Relatively high levels of biotransformation enzymes and relatively high oxygenation

Answer

Relatively high levels of biotransformation enzymes and relatively low oxygenation

---

This page titled 2.6.1.5: Patterns of Toxic Injury is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **1.6: Patterns of Toxic Injury** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.1.6: Final Evaluation

1. General rule of thumb for apoptosis is that it is more commonly seen at higher levels of toxicant exposure while necrosis occurs more frequently at relatively lower levels of toxicant exposure. *True or False?*

True

False

Answer

False

2. Neoantigen formation results when a xenobiotic or its metabolite binds to a larger protein to form a novel molecule that elicits an immune response. Molecules that trigger this immune response are called:

Haptens

Free radicals

Electrophils

Poisons

Answer

Haptens

3. Lipid peroxidation occurs when an _____ compound steals an electron from a membrane phospholipids resulting in the production of a fatty acid radical.

Electrophobic

Hydroxyl radical

Electrophilic

Hydrophilic

Answer

Electrophilic

4. Electron transfer can result in oxidation of some endogenous macromolecules and formation of _____ such as superoxide ion (•O2-) and hydroxyl radical (HO•).

Enzymes

Hormones

Free radicals

Proteins

Answer

Free radicals

5. Upon exposure of human skin with leaves of poison ivy, the toxin found in this plant binds to membranes on skin cells to stimulate an immune response resulting in blistering rash. This toxin is known as:

Urushiol

Ricin

Tetrodotoxin

Pyrethrins

Answer

Urushiol

6. Which of the following covalently binds to the acetaminophen metabolite N-Acetyl-P-Benzoquinone Imine (NAPQI) to detoxify it?

N-acetylcysteine

Superoxide dismutase

Catalase

Amylase

Answer

N-acetylcysteine

7. Differences between Necrosis and Apoptosis include the following EXCEPT:

Necrosis is a degenerative process while Apoptosis is an active process

Necrosis triggers inflammatory response while Apoptosis does not incite inflammatory response

Necrosis results in loss of energy while Apoptosis does not result in loss of energy

Necrosis is an active process while Apoptosis is a degenerative process

Answer

Necrosis is an active process while Apoptosis is a degenerative process

8. A regenerative response that results in an increase in cell size is termed _____.

Hypertrophy

Hypoplasia

Hyperplasia

Atrophy

Answer

Hypertrophy

9. _____ helps to terminate lipid peroxidation when present in sufficient quantities.

Vitamin E

Hydroxyl radical

Vitamin A

Oxygen

Answer

Vitamin E

10. An “adduct” is the product of an irreversible bond between the toxicant and target molecule; this type of binding is known as:

Noncovalent binding

Covalent binding

Hydrogen abstraction

Ionic binding

Answer

Covalent binding

11. An example of decreasing the reactivity of a target site to a toxicant is describe as:

Binding of toxicant to target receptors

Downregulation of toxicant target receptors

Increased metabolism of alcohol within the stomach and liver

Stimulation of membrane receptors

Answer

Downregulation of toxicant target receptors

12. 10% of acetaminophen bioactivated by Cytochrome P450 in the liver to a toxic metabolite is called:

Superoxide dismutase

N-acetylcysteine

N-acetyl-P-Benzoquinone Imine (NAPQ1)

Superoxide ion

Answer

N-acetyl-P-Benzoquinone Imine (NAPQ1)

This page titled 2.6.1.6: Final Evaluation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **Final Evaluation** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.2: Biochemistry and Molecular Genetics

Biochemistry is the study of chemical processes within and related to living organisms. In this e-module you will learn about biomolecules and cell components, cell structure and subcellular compartments, DNA and RNA metabolism, and epigenetic mechanisms.

2.6.2.1: Introduction to Biomolecules and Cell Components

2.6.2.2: Cell Structure and Subcellular Compartments

2.6.2.3: DNA and RNA Metabolism

2.6.2.4: Epigenetic Mechanisms

2.6.2.5: Section 2 Final Evaluation

This page titled 2.6.2: Biochemistry and Molecular Genetics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.2.1: Introduction to Biomolecules and Cell Components

Learning Objectives

- 1: Define the basic structure of biomolecules, such as: amino acids and proteins, carbohydrates, fatty acids, triacylglycerol, phospholipids, steroids and nucleic acids.
- 2: Define the meaning and significance of essential and non-essential amino acids.
- 3: Understand the function of enzymes.
- 4: Define the basic structure of ribonucleic acid (RNA) and deoxyribonucleic acid (DNA).

1.1

Amino Acids and Proteins

Amino acids are the basic units of proteins. All amino acids present in proteins carry a carboxyl- and an amino group, hydrogen and variable side chains (R) at a single α – carbon atom.

Amino Acid Basic Structure:

Every amino acid has four components linked together with a central carbon atom α – carbon:

- **Amino group (NH2)**
- **Carboxylic acid group (COOH)**
- **Hydrogen atom (H)**
- **R-group, which varies with each amino acid (R)**

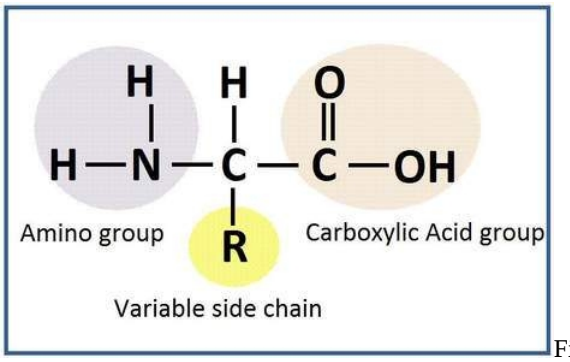


Figure 2.6.2.1.1: Basic structure of an amino acid. Credit: Aline de Conti

R groups may be:

- Hydrophobic
- Hydrophilic
- Charged R-groups: *positive or negative charged*
- Special R-groups: *conjugated with other molecules*

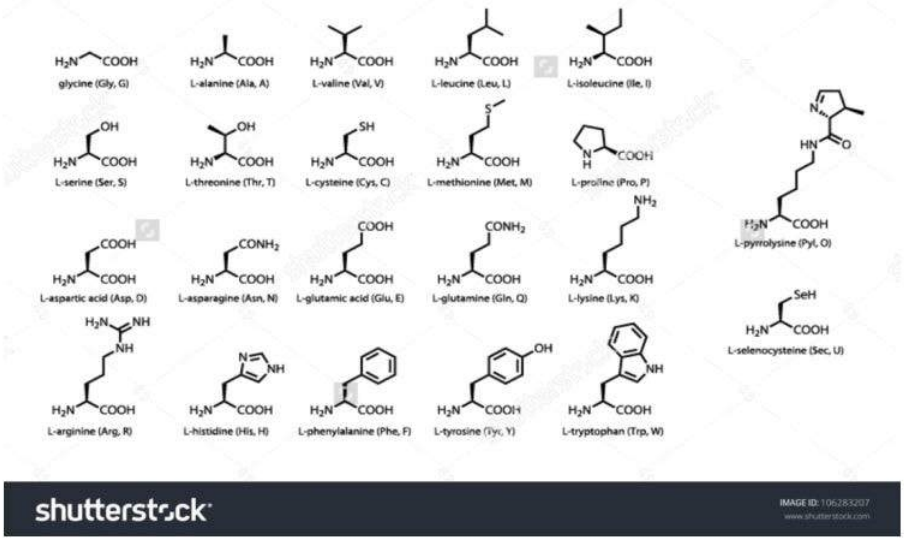
Amino Acids are classified into two groups.

Essential: Humans cannot synthesize them and must be obtained directly from food (phenylalanine, valine, threonine, tryptophan, methionine, leucine, isoleucine, lysine, histidine, cysteine, and arginine).



Non-essential: The human body is able to produce them (glycine, alanine, serine, asparagine, glutanine, tyrosine, aspartic acid, glutanic acid, and proline).

Below, you can see different structures of the most common amino acids in humans. Amino acids link together, in a reaction known as **peptide bond**, to form proteins.



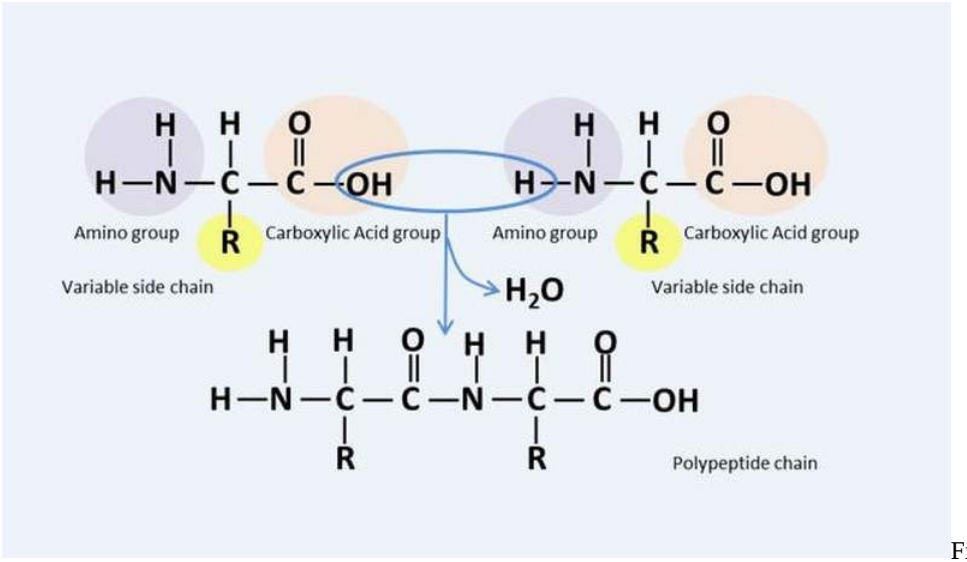


Figure 2.6.2.1.2: Peptide bond. Credit: Aline de Conti

Levels of Protein Structure

Primary (1°) Structure

The sequence of amino acids in a protein is named as **primary structure.** The amino acids are linked via peptide bonds formed with the carboxylic acid group of one amino acid and the amino group of a other amino acid.

Secondary (2°) Structure

The secondary structure is the way a polypeptide folds to form **α-helix**, **β-strand**, or **β-turn**.

Tertiary (3°) Structure

The tertiary structure is the way the polypeptide chain coils and turns to form a complex molecular shape. Additionally, tertiary structure starts to develop an active sites of proteins where critical actions and interactions will take place.

Quaternary (4°) Structure

The quaternary structure is the combination of the multiple protein subunits that interact to form a single, larger, biologically active protein.

Protein Functions

Proteins have several functions in the human body including hormonal, enzymes, structural proteins in cell membranes, proteins also receive signals from outside the cell and mobilize intracellular response, and they are part of the immune system.

Enzymes are specialized proteins that accelerate a chemical reaction by serving as a biological catalyst. By catalyzing these reactions, enzymes cause them to take place one million or more times faster than in their absence. Several biochemical reactions important for cellular maintenance occur due to enzymes activity. For example: environmental response and metabolic pathways.

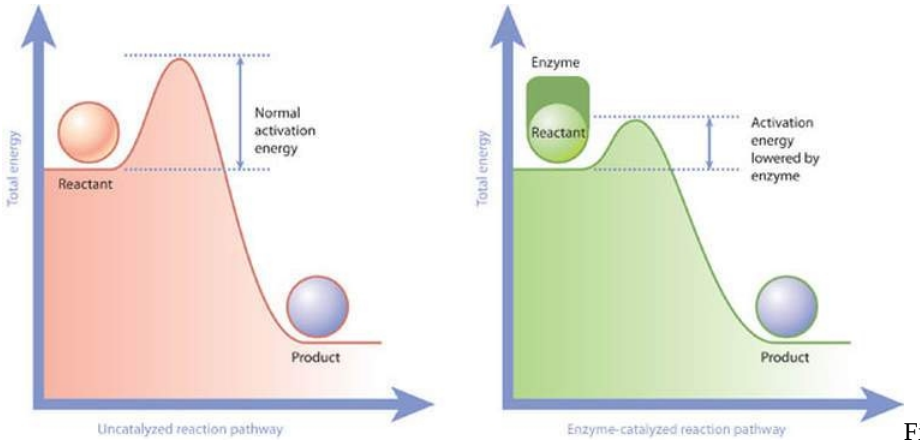
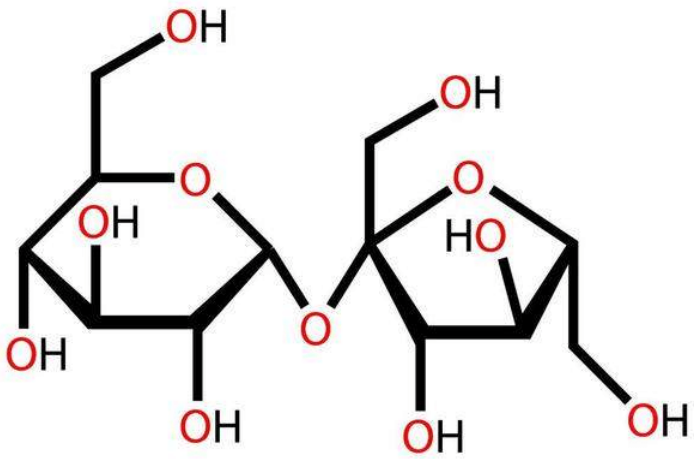


Figure 2.6.2.1.3: Enzymes and Activation Energy - © 2010 Nature Education

1.2

Carbohydrates



Carbohydrates are made of molecules of carbon (C), Hydrogen (H) and Oxygen (O), and are composed of recurring monomers called monosaccharides (which typically form ring structures). A common name of monomers and dimers is ‘sugar’.

Carbohydrates are classified into three subtypes.



1. Monosaccharides: 1 unit of monomer. Examples: fructose, glucose, galactose.

Present in fruits, etc.

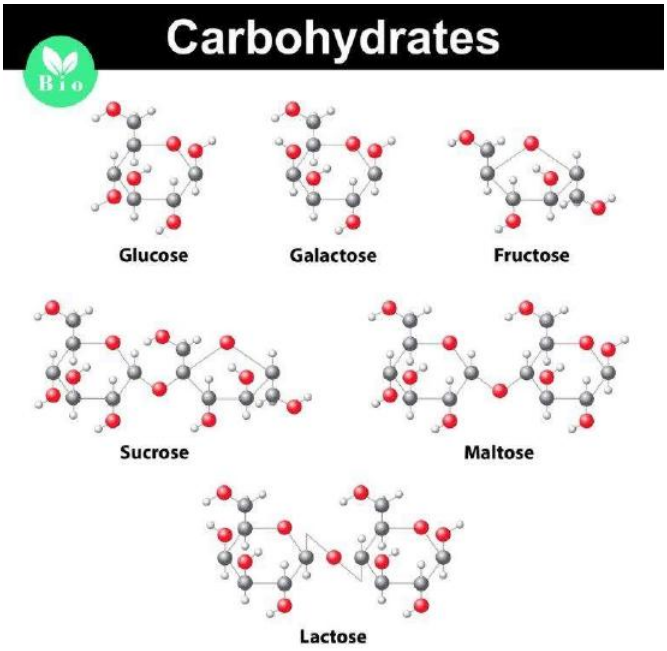
2. Disaccharides: 2 units of monosaccharides. Examples: lactose, maltose, sucrose.

Present in milk, etc.

3. Polysaccharides: Many monosaccharide units. Examples: cellulose, glycogen, starch.

Present in breads, grass, etc.

Carbohydrates are a group of macromolecules that are important energy source required for various metabolic activities. Carbohydrates may bind to proteins and lipids that play important roles in cell interactions e.g. receptor molecules and immune system e.g. antigens.



1.3

Lipids

Lipid molecules are mainly hydrophobic molecules i.e. found in areas away from water molecules, but also present smaller hydrophilic parts that are important for its biological function. The major roles of lipid molecules are to serve as storage of biological energy (Example: **triacylglycerols**) and provide the building blocks for biological membranes (Example: **phospholipids** and **cholesterol**). Although there are other types of lipids, in this topic we will discuss the structure and function of these main groups of lipids.

Triacylglycerols

Triacylglycerols are composed of fatty acids and glycerol.

- **Glycerol** is a simple three-carbon molecule with hydroxyl groups at each carbon.
- **Fatty acids** are chains of carbon molecules with a carboxylic acid (COOH) in the first carbon and a CH3 (methyl) group at the end of the chain.



Fatty acids can be...

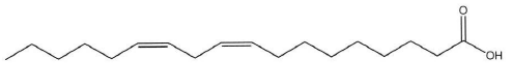
Saturated: Fatty acids contain only single carbon-carbon bonds, and all of the carbon molecules are bonded to the maximum number of hydrogen molecules.

Unsaturated: Fatty acids have at least one double carbon-carbon bond with the potential for additional hydrogen atom bonding still existing for some of the carbon atoms in the backbone chain. If more than one double bond is present, the term polyunsaturated is used.

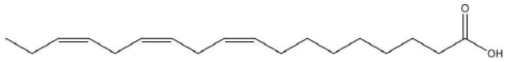
Essential Fatty Acids

Examples of two essential fatty acids, linoleic acid (known as **omega-6;ω-6**) and linolenic acid (known as **omega-3; ω-3**). These fatty acids present double bonds at the sixth and third carbon atoms, respectively, counting from the methyl end of their chains. They are considered essential because humans do not have the ability to produce double bonds at these locations and, therefore, must obtain these two fatty acid from vegetable oils.

linoleic acid (omega-6 fatty acid)



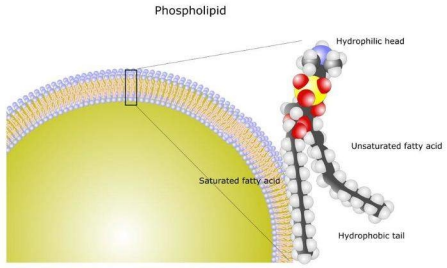
alpha-linolenic acid (omega-3 fatty acid)



Phospholipids

Phospholipids are the major component of cell membranes. They form lipid bilayers because of their amphiphilic characteristic.

The structure of the **phospholipid** molecule generally consists of two hydrophobic fatty acid "tails" and a hydrophilic "head" consisting of a phosphate group (PO_4^{-3}) attached to the third glycerol carbon. This head group is usually charged, creating a part of the lipid that is hydrophilic, and wants to be near water, a quality that is essential for the formation of biological membranes and many lipid functions.



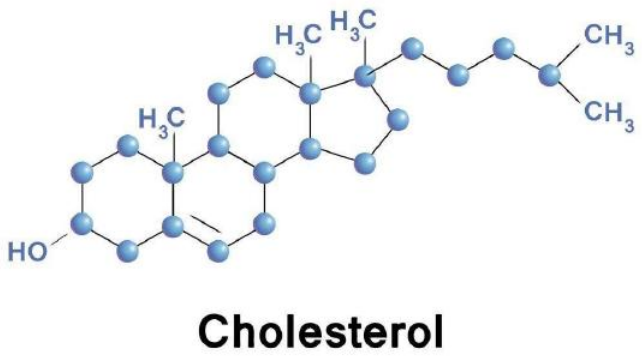
Steroids

Steroids are lipids that have four rings made of carbon atoms—three rings have six sides and one has five sides—with a six-carbon ring tail. Examples: bile salts, cholesterol, the sexual hormones estrogen, progesterone and testosterone, corticosteroids and pro-vitamin D.

Cholesterol

Cholesterol is an important molecule found only in eukaryotic organisms with a variety of functions. Cholesterol is also a component of biological membranes and its main function is to control the fluidity of membranes. Cholesterol does not like to be exposed to water environments, preferring to be shielded by other hydrophobic molecules such as lipids or hydrophobic parts of proteins.

Cholesterol also serves as the primary source for the production of steroid hormones, bile salts, and vitamin D.

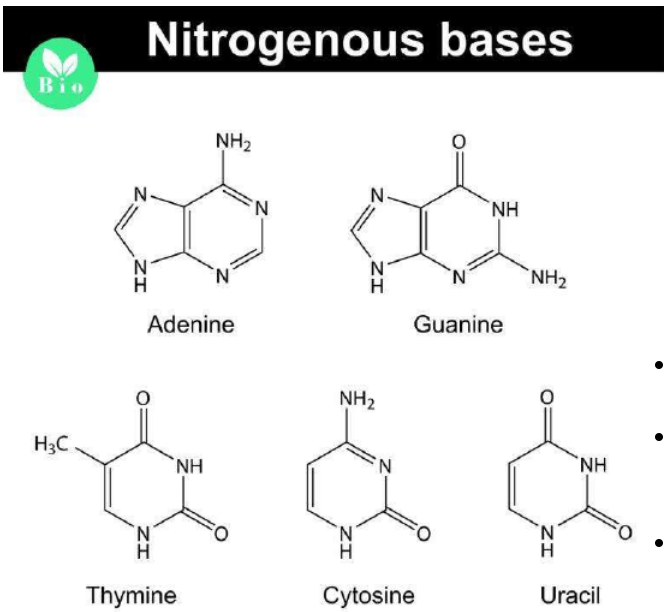


Cholesterol

1.4

Nucleosides and nucleotides are involved in the preservation and transmission of the genetic information of all living creatures. In addition, they play roles in biological energy storage and transmission, signaling and regulation of various aspects of metabolism.

These molecules can be divided into two major families.



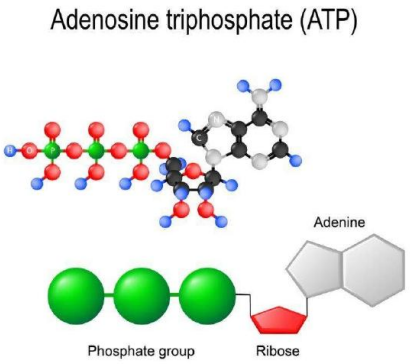
Purines: They are two-ring structures: adenine and guanine.

Pyrimidines: They are one-ring structures: thymine, cytosine, and uracil.

The unique structure and interaction of these molecules serve as the fundamental building block of RNA and DNA molecules and allow fundamental processes of DNA replication and protein synthesis to occur.

Components of Nucleotides

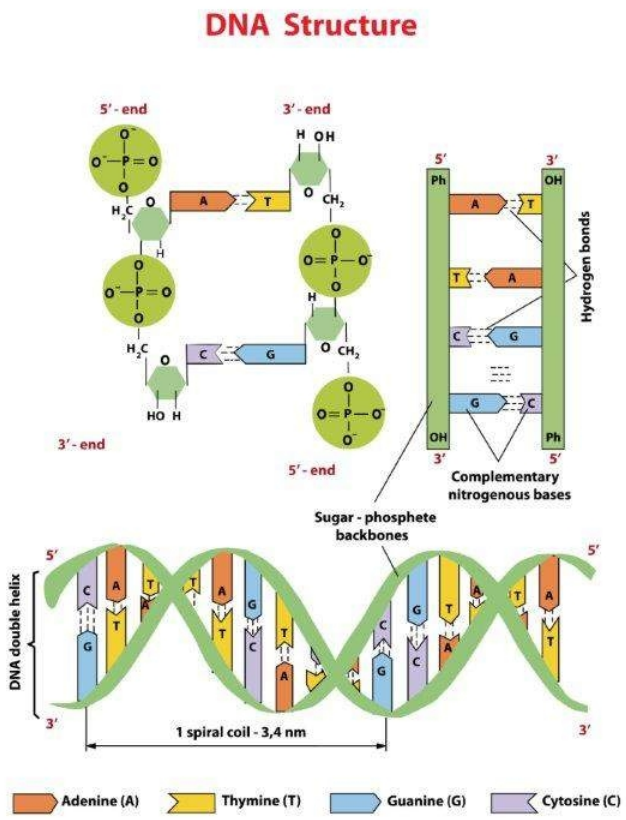
- **Nitrogenous base:** The nitrogenous base of a nucleoside or nucleotide may be either a purine or a pyrimidine.
- **Carbohydrate:** The carbohydrate component of nucleosides and nucleotides is usually the sugar ribose for RNA molecules and deoxyribose for DNA molecule
- **Phosphate Group:** One or more phosphate groups (PO4−3) may be attached to the carbon 5 of the carbohydrate molecule.



1.5

DNA

DNA stands for deoxyribonucleic acid.
It is an extremely long molecule that forms a double-helix.



DNA components:

- Sugars - Deoxyribose

- Phosphates - (PO4−3)

- Base - cytosine (C), guanine (G), adenine (A) and thymine (T).

The DNA consists of two strands attached to each other by hydrogen bond created by nucleotide pairing (A-T and C-G).

The **double-helix** structure of DNA is important for its function because these two bonded strands can temporarily separate to allow for DNA replication.

The sequences of nucleotides (A, C, T, G) in the DNA molecule will make up the **genes** and, subsequently, proteins are referred to as “expressed sequences” or “exons.” Sequences that do not code for a protein are called “intervening sequences” or “introns.”

Human Genome

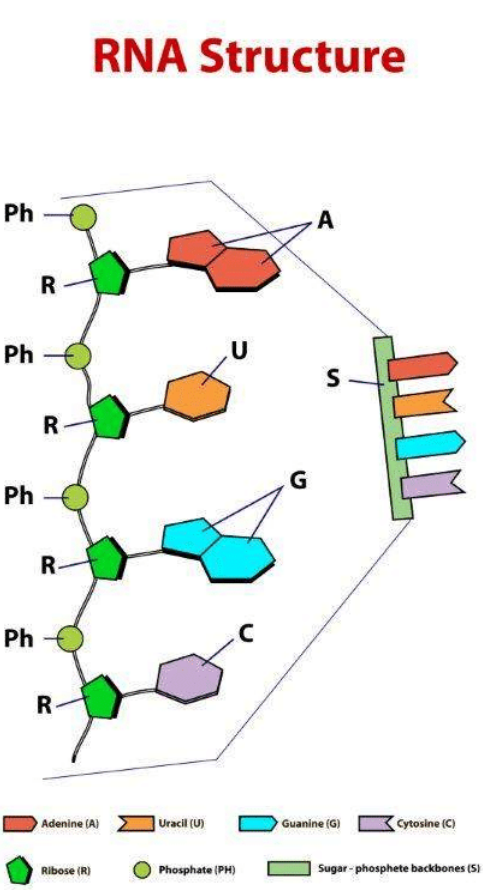
The genome of humans is estimated to contain approximately 20,000–25,000 different genes arranged on multiple chromosomes.

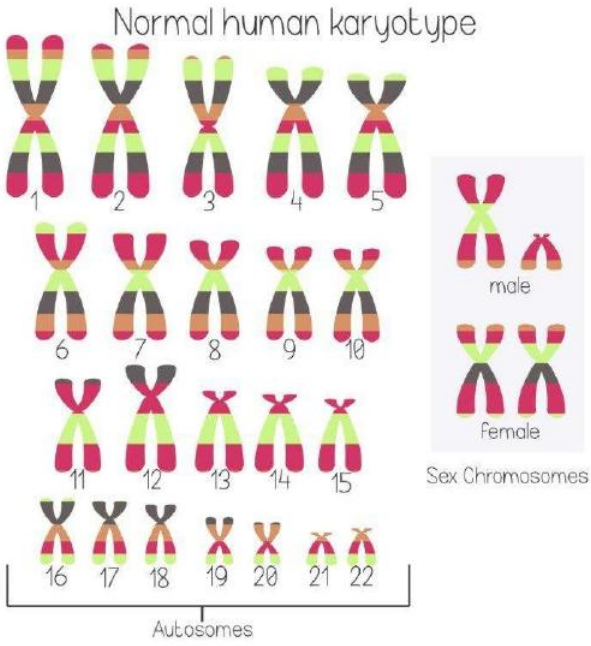
Twenty three pairs of chromosomes:
Twenty two pairs (autosomes).
One pair (sex chromosome) (xx) (female) or (xy) (male).

Humans have 23 pairs of chromosome in every cell (except mature red blood cells); Gametes or sex cells (sperm and eggs) have half the normal complement of chromosomes.

1.6

RNA





RNA stands for ribonucleic acid. RNA molecules are single strands.

RNA components:

- Sugars - Ribose
- Phosphates - (PO_4^{-3})
- Base: cytosine (C), guanine (G), adenine (A) and uracil (U)

RNA molecules often form secondary (2°) structures and may interact with DNA, other RNA molecules, and proteins. These interactions help to define the particular function of each type of RNA.

Types of RNA molecules and functions:

Messenger RNA (mRNA):

Molecules which function as the transmitter of genetic information from the DNA genetic code to the resulting protein.

Transfer RNA (tRNA)

Molecules that carry amino acids and match them with a specific mRNA sequence during protein synthesis.

Ribosomal RNA (rRNA)

Molecules associated with proteins and are responsible for the synthesis of protein molecules.

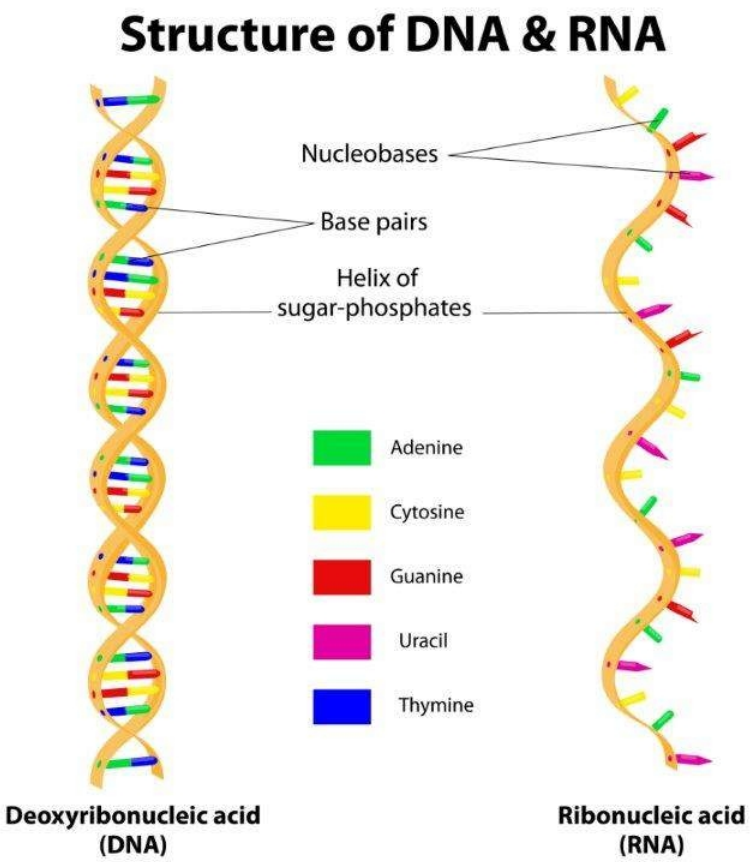
Regulatory RNA

Molecules involved in regulation of DNA expression, posttranscriptional mRNA processing, and the activity of the transcribed mRNA message.

The basic structure of DNA and RNA are similar, however with 3 main differences:

Nitrogenous Bases: Three of the nitrogenous bases are the same in the DNA and RNA: adenine, cytosine, and guanine. The fourth base for DNA is thymine while for RNA it is uracil. Thymine and uracil both bind to adenine.

Number of Strands: The DNA molecule is usually double-stranded and most cellular RNA molecules are single-stranded.



Type of Sugar: In the DNA molecule the sugar is **deoxyribose** and in the RNA molecule the sugar is **ribose**.

Topic 1: Key Points

In this section, we explored the following main points:

- 1: Amino acids link together, in a reaction known as peptide bond, to form proteins.
- 2: One important function of protein is to act as enzymes to accelerate chemical reactions.
- 3: Carbohydrates are important energy source required for various metabolic activities and may bind to proteins and lipids that play important roles in cell interactions
- 4: Lipid molecules serve as storage of biological energy and provide the building blocks for biological membranes
- 5: DNA and RNA structures have 3 main differences .The nitrogenous bases (DNA has thymine and RNA has uracil). The DNA molecule is usually double stranded and most of the RNA molecules are single stranded. In the DNA molecule the sugar is deoxyribose and in the RNA molecule the sugar is ribose.

Knowledge Check

1. What type of nucleic acid does thymine belong to?

Answer

DNA

2. Uracil?

Answer

RNA

3. Enzymes are...

specialized proteins that accelerate a chemical reaction by serving as a biological catalyst.

specialized proteins that stops a chemical reaction.

Answer

specialized proteins that accelerate a chemical reaction by serving as a biological catalyst.

4. **A nucleotide consists of...**

Check all that apply.

- A sugar (either deoxyribose or ribose)
- Uracil as the nitrogen base
- A phosphate group
- One of the four nitrogen bases

Answer

A sugar, A phosphate group, and One of the four nitrogen bases

5. **Lipid molecules are mainly hydrophilic molecules.** *True or False?*

Answer

false

This page titled 2.6.2.1: Introduction to Biomolecules and Cell Components is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **2.1: Introduction to Biomolecules and Cell Components** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.2.2: Cell Structure and Subcellular Compartments

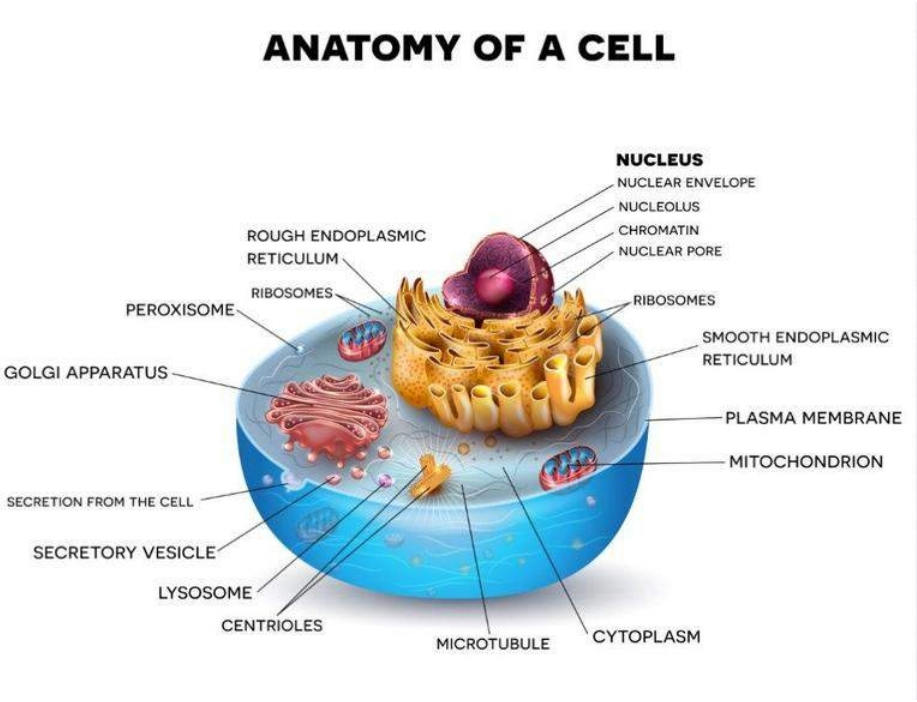
Learning Objectives

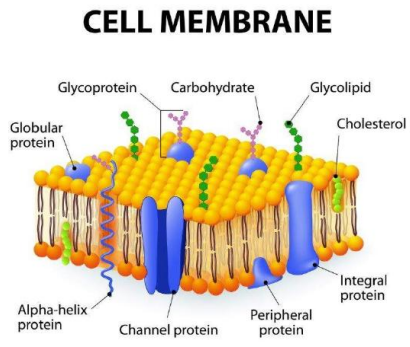
- 1: Describe the basic components of cell structure and organelles.
- 2: Understand the chromatin organization into nucleus.
- 3: Understand basic mechanisms of cell signaling to respond to changes in the environment.

2.1

Structure of a cell

Cells are the smallest unit of life. A typical eukaryotic cell and it components are illustrated in the next figure.

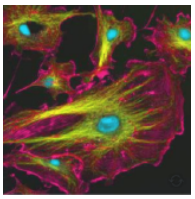




Membrane: Plasma membrane, nuclear envelope, and membranes provide separation of different environments to permit a variety of biological functions. Membranes are dynamic and fluid structures that allow selective movement of ions, energy sources, vitamins and cofactors, and waste. Membrane components include lipid, carbohydrates and protein molecules. Membrane fluidity is very important for its functions and, as a consequence, is important in disease processes, as well as, treatments.Figure 2.6.2.2.1: Structure of cell membrane

Membranes are composed of lipids arranged in a **lipid bilayer,** with the hydrophilic glycerol and phosphate “head” groups of the lipid molecules forming the two outside layers and the hydrophobic “tail” groups arranged inside. Proteins are the second major part of biological membranes and makeup approximately 20%–80% of both the structural and functional components of these membranes. Many of these proteins are embedded into the membrane and stick out on both sides; these are called **transmembrane proteins.**

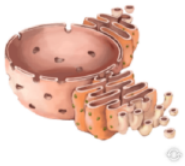
Subcellular compartments



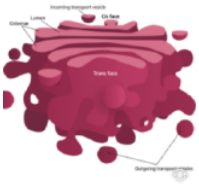
CYTOSKELETON

is a structure that helps cells maintain their shape and internal organization, and it also provides mechanical support that enables cells to carry out essential functions like division and movement. The cytoskeleton is made up of microtubules, actin filaments, and intermediate filaments.

ENDOPLASMIC RETICULUM (ER)

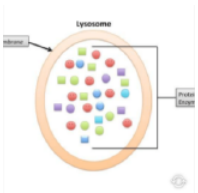


It consists of a system of sac- and tube-like structures, which locally expand into cisterns. Its internal lumen is connected with the intermembrane space of the nuclear membrane. Part of the ER is studded on the outside with ribosomes (rough ER), which take part in protein synthesis. The other part of the ER is free of ribosomes (smooth ER). Enzymes of the smooth ER are involved in the synthesis of fatty acids. The smooth ER also plays a role in detoxification by hydroxylation reactions.



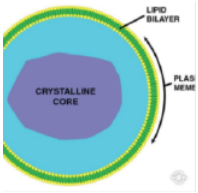
GOLGI APPARATUS

This organelle consists of stacks of flattened membrane sacs. Their main function is the further processing and sorting of proteins and their export to the final targets. In most cases, these are secreted as membrane proteins. In addition, the Golgi apparatus also produces polysaccharides.



LYSOSOMES

Lysosomes are vesicles enclosed by a lipid bilayer. These organelles are filled with many enzymes for polysaccharide, lipid, protein and nucleic acid degradation. They act also on intracellular material to be removed and even contribute to the apoptosis of their own cell. Lysosomes of special cells (e.g., macrophages) destroy bacteria or viruses as a defense mechanism.



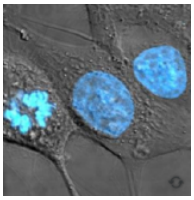
PEROXISOMES

Peroxisomes are surrounded by a single membrane. They are generated from components of the cytosol and do not bud from other membranes. The main task of these organelles is the performance of monooxygenase (hydroxylase) or oxidase reactions, which produce hydrogen peroxide (H2O2).



MITOCHONDRIA

In a typical eukaryotic cell, there are in the order of 2000 of these organelles, which are often of ellipsoidal shape. They have a smooth outer membrane and a highly folded inner membrane with numerous invaginations (cristae), which contain most of the membrane-bound enzymes of mitochondrial metabolism. Mitochondria are the site of respiration and ATP synthesis, but also of many other central reactions of metabolism, e.g., citrate cycle, fatty acid oxidation, glutamine formation, and part of the pathway leading to steroid hormones. Mitochondria are the only organelles which are equipped with their own (circular) DNA, RNA and ribosomes and thus can perform their own protein synthesis.



NUCLEUS

All eukaryotic cells show the presence of a separate nucleus, which contains the major portion of the genetic material of the cell (DNA). The nuclear DNA is organized in a number of chromosomes. The nucleus is surrounded by a double membrane of lipid bilayers with integrated proteins, called the nuclear membrane (also known as the nuclear envelope). Nuclear pores span the nuclear membrane and enable the transport of proteins, rRNA etc.

Image credits

2. DataBase Center for Life Science (DBCLS) licensed under CC BY 4.0

3. Kelvin Song licensed under CC BY 3.0

4. Lumoreno licensed under CC BY-SA 3.0

5. Agateller licensed under CC BY-SA 3.0

2.3

Structure of Chromatin

Chromosomal DNA is packaged inside microscopic nuclei by its association of histones H2A, H2B, H3, and H4. These are positively-charged proteins that strongly adhere to negatively-charged DNA and form complexes called nucleosomes. (11-nm "beads on a string" structure). Each nucleosome is composed of DNA wound 1.65 times around eight histone proteins. Nucleosomes fold up to form a 30-nanometer chromatin fiber, which forms loops averaging 300 nanometers in length. The 300 nm fibers are compressed and folded to produce a 250 nm-wide fiber, which is tightly coiled into the chromatid of a chromosome. The level of structure varies depending on the cell cycle stage and, as a result, the requirement for DNA transcription or replication.

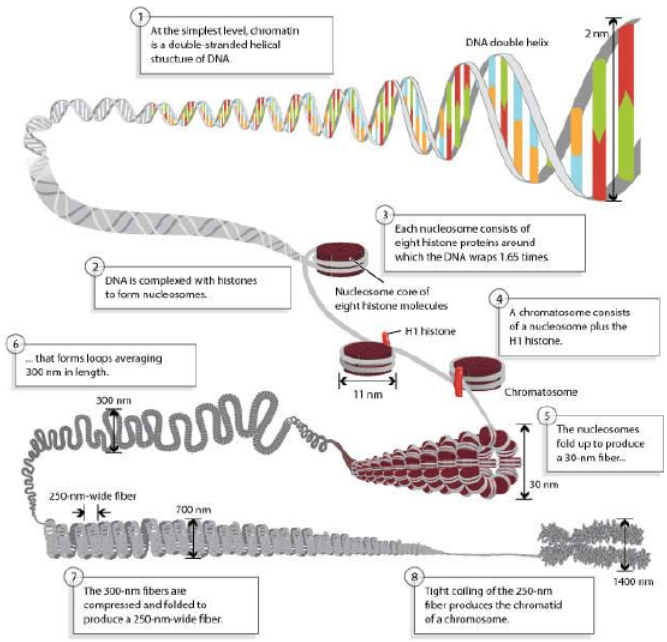
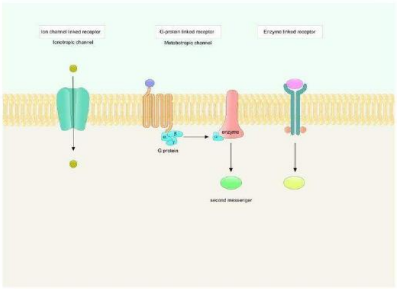


Figure 2.6.2.2.2: Chromosomes are composed of DNA tightly-wound around histones - © 2010 Nature Education

Chromatin: The most important structure inside the nucleus is **chromatin**, consisting, in humans, of the 46 chromosomes. Chromatin is the combination or complex of DNA and proteins that make up the contents of the nucleus of a cell. The most abundant protein in the nucleus are histones. Histones are rich in basic amino acids (positively charged), which interact with negative charges of the DNA.

2.4

Cell Signaling



Cell signaling consists in the ability of cells to respond to environment changes through signals received outside their borders. Cells may receive many signals simultaneously and also send out messages to other cells.

Cells have proteins called receptors, that are generally transmembrane proteins, which bind to signaling molecules outside the cell and subsequently transmit the signal through a sequence of molecular switches to internal signaling pathways and initiate a physiological response. Different receptors are specific for different molecules. In fact, there are hundreds of receptor types found in cells, and varying cell types have different populations of receptors.

Figure 2.6.2.2.3: Main types of transmembrane receptors

Examples of receptors membrane:

- G-protein-coupled receptors
- Ion channel receptors
- Enzyme-linked receptors

Receptor may be located in the cellular membrane (important to receive extracellular signals) but also may be present inside the cell or inside the nucleus.

Topic 2: Key Points

In this section, we explored the following main points:

- 1: The characteristics of the cell membrane, such as lipid bilayer and transmembrane proteins are very important for the cell function.
- 2: The chromatin is the combination of complex DNA and proteins that make up the contents of the nucleus of a cell.
- 3: Cells have proteins called receptors that are important to receive extracellular signals and initiate signaling pathways in the cell.

Knowledge Check

1. Histones are proteins present in the nucleus of the cells with the function of:

interacting with the DNA to form nucleosomes.

regulating the fluidity of cell membranes.

Answer

interacting with the DNA to form nucleosomes.

2. Membranes are composed of lipids arranged in a ...

Check all that apply.

- lipid monolayer, with...
- lipid bilayer, with...
- ...the hydrophilic glycerol and phosphate “head” groups of the lipid molecules forming the two outside layers and the hydrophobic “tail” groups arranged inside.
- ...the hydrophilic and hydrophobic “tail” groups arranged inside.

Answer

lipid bilayer, with...the hydrophilic glycerol and phosphate “head” groups of the lipid molecules forming the two outside layers and the hydrophobic “tail” groups arranged inside.

This page titled 2.6.2.2: Cell Structure and Subcellular Compartments is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **2.2: Cell Structure and Subcellular Compartments** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.2.3: DNA and RNA Metabolism

Learning Objectives

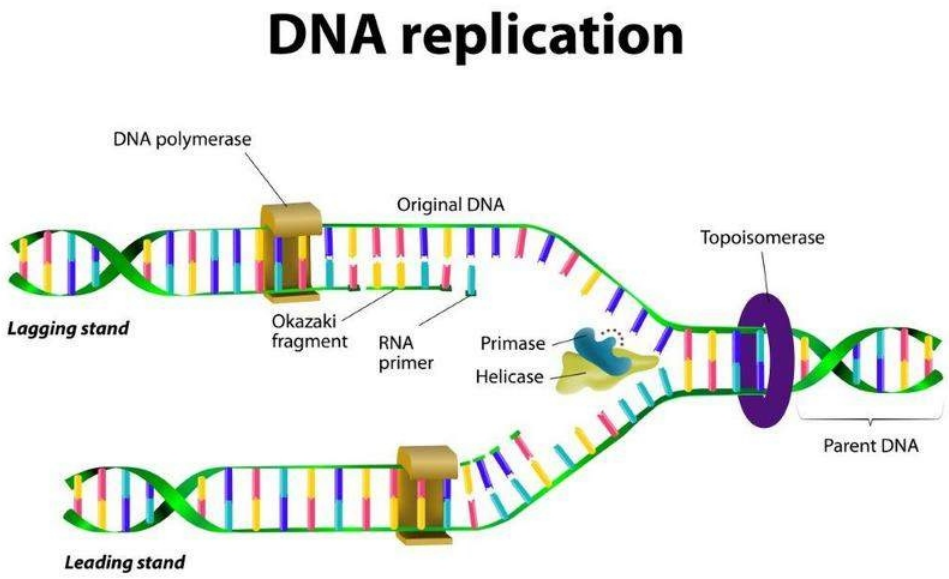
After completing this lesson, you will be able to:

- Understand the mechanism of DNA replication.
- Understand the process of gene expression.
- List the main types of DNA mutation and mechanism of DNA repair.

DNA Replication

The process of DNA replication consists of uncoiling double-stranded DNA, copying each DNA strand and then separating the two, new, double-stranded copies. The process starts at an **origin of replication (ori)**, a nucleic acids sequence where the replication can start. There are around 100,000 origins of replication in each human cell. This means that the DNA replication may start simultaneously in different positions at the same time. The **replication fork** is the point where two DNA strands, one termed the **leading strand** and the other the **lagging strand**, are separated and DNA copying occurs. The coiled-coil, double-helical DNA structure is initially unwound by the enzyme **DNA helicase** by breaking the hydrogen bonds between complementary nucleic acids. **Single-stranded binding proteins** attach to the new DNA strands to keep them separated.

An enzyme termed **primase** then produces a short strand of RNA to serve as a primer for the remainder of the process. The enzyme **DNA polymerase** replicates each DNA strand in the 5′ to 3′ direction by adding the correct, matching nucleotide triphosphate to the 3′-hydroxyl end of the primer strand. As each new nucleic acid is added, a new phosphodiester bond is formed, utilizing the energy contained in the remaining diphosphate group.



RNA Transcription

RNA Transcription is the process whereby a particular segment of DNA is transcripted into an equivalent RNA sequence.

- mRNA: For genes that codes for a protein.
- tRNA: For atransfer RNA.
- rRNA: For assembly of a ribosome.
- miRNA (micro RNA): Which binds to mRNA andinhibits its translation.
- siRNA (small-interfering RNA): Which binds to mRNA and aids in its degradation.
- snRNA (small nuclear RNA): Which participates in RNA processing as part of the spliceosomes.

- snoRNA (small nucleolar RNA): Which participate in nucleolar RNA processing.

The transcription starts with binding of the enzyme **RNA polymerase** to a **promoter** sequence on the DNA, a regulatory region that dictates where the transcription should start.

The DNA is transcribed from 3′ to 5′ and occurs only on one of the DNA strands, the **template strand.**

As in DNA replication, energy for the formation of the phosphodiester bond is derived from hydrolysis of the two terminal phosphate bonds of the nucleoside triphosphate.

Multiple RNA polymerases can transcribe on a single DNA gene sequence, allowing rapid production of the RNA product.

Enhancer: This is a short region of DNA that can be bound by transcription factors to increase or facilitate the transcription of a particular gene. They can be located up away from the gene, upstream, or downstream from the start site.

Transcription Factors (TFs): These include a wide number of proteins, excluding RNA polymerase, that promotes (as an activator), or blocks (as a repressor) the recruitment of RNA polymerase to DNA. TFs bind to promoter regions of DNA.

Promoter Region: These are specific DNA sequences, usually located upstream and near the transcription start sites of genes and serves as a binding site for proteins called transcription factors that recruit RNA polemerase. Example: TATA bos, CpG island.

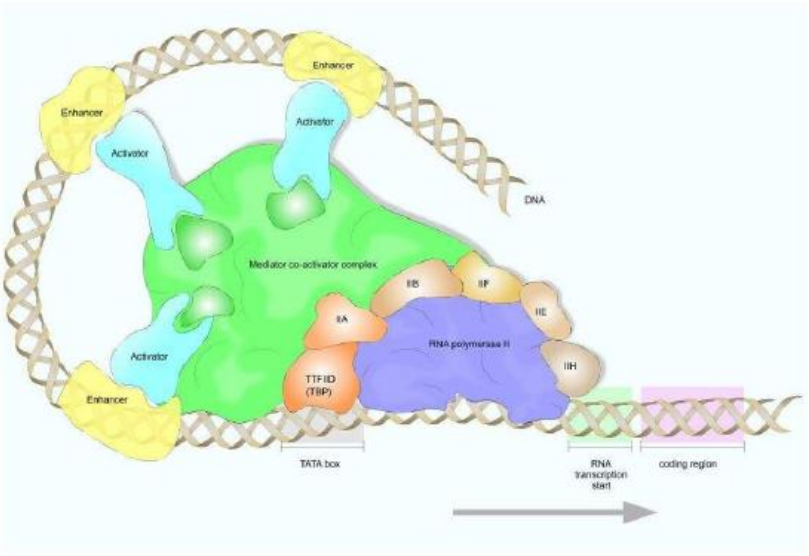


Figure 2.6.2.3.1: Machinery of RNA transcription

RNA Processing

The newly synthesized RNA transcripts are processed prior to their use in the cell as mature RNA.

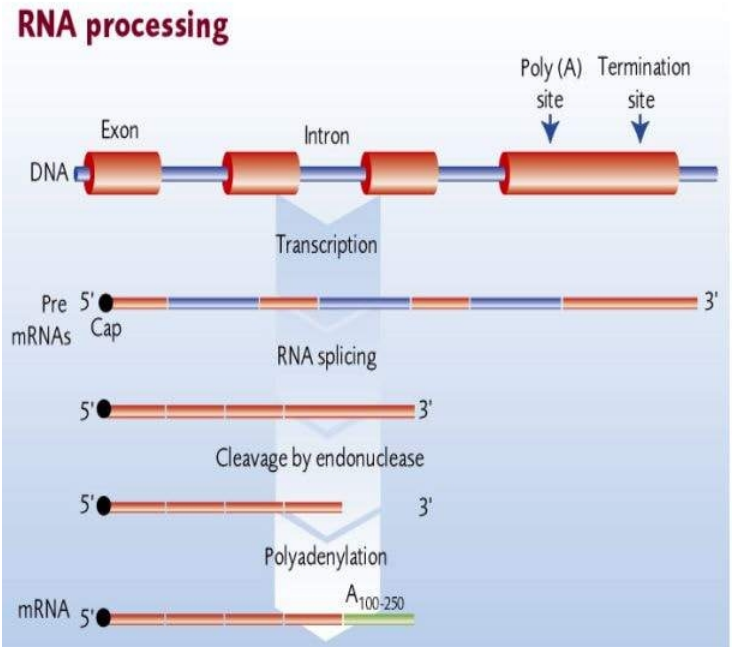


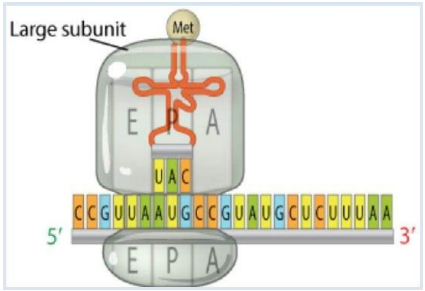
Figure 2.6.2.3.2: RNA processing

A 7-methyl guanosine nucleic acid is added to the 5′-end (known as a **5′ cap**) of the pre-mRNA as it emerges from RNA polymerase II (Pol II). The cap protects the RNA from being degraded by enzymes and serves as an assembly point for the proteins to begin translation to protein.

Removal of **introns** present in the pre-mRNA and splicing of the remaining **exons**, in a process called **RNA splicing**. The continuous series of DNA bases coding for a protein are interrupted by base sequences that are not translated. The translated sequences are referred to as exons (expressed sequences) and the nontranslated sequences as introns (intervening sequences). This completes the mRNA molecule, which is now ready for export to the cytosol. (The remainder of the transcript is degraded, and the RNA polymerase leaves the DNA.)

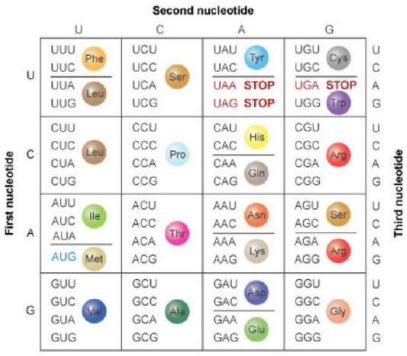
3.4 Protein Translation

Protein synthesis requires the interaction of **mRNA, tRNA**, several accessory proteins, called **initiation factor (IF)** and **elongation (EF)** factor, and **ribosomes.**



The large ribosomal subunit binds to the small ribosomal subunit to complete the initiation complex. The tRNA molecule bind to one amino acids at the top of the structure. In the base if the tRNA there is the three base sequence known as the **anticodon**. This anticodon binds, through base pairing, to a three base codon on mRNA. It is this interaction between the mRNA and the amino acid-tRNA that provides the high degree of fidelity observed in the transfer of genetic information from DNA to proteins.Figure 2.6.2.3.3: Transfer RNA (tRNA) - © 2013 Nature Education

The Triplet Code (Genetic Code)



A sequence of three bases in DNA identifies each of the 20 amino acids that are to be incorporated into the newly synthesized protein. This information is incorporated into mRNA which is synthesized using DNA as the template. Considering that three bases are required at a minimum, and we have 4 nucleotides (43 = 64 code words are possible). This is more than the 20 amino acids. In fact, many triplets are used to define one same amino acid. In addition, some “extra” triplet sequences are used as stop codons to terminate protein synthesis. AUG is used as the start codon for the N-terminal amino acid in eukaryotes.

Using the same binding rules as DNA double strands (e.g., a tRNA that binds the starting mRNA codon AUG has an anticodon sequence of UAC), insuring the specific order of AAs required for proper production of the protein.

Figure 2.6.2.3.4: Genetic code - © 2014 Nature Education

RNA plays three distinct and important roles:

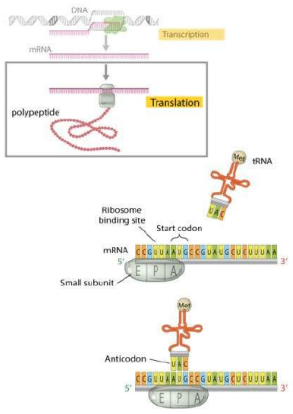
- mRNA: The intermediary between gene and protein – provides the message.
- tRNA: The key or adaptor – reads the genetic code, brings amino acids to the growing polypeptide chain
- rRNA: In ribosome, provides a scaffold for protein synthesis, catalyzes peptide bond formation

The IF and EF accessory proteins serve a number of roles, including enabling binding of the mRNA molecule to the ribosome, movement of the mRNA along the ribosome to the start point of the synthesis, docking of the tRNA–amino acid, and movement of the mRNA and growing peptide chain, as well as accuracy assurance.

The protein biosynthesis, can be divided into three phases: initiation, elongation, and termination.

Initiation

Initiation of protein synthesis begins when the protein initiator factor IF-3 binds to the small subunit of the ribosome and causes its dissociation. The small ribosomal subunit then binds to the 5’ side of mRNA which carries information in a triplet code from DNA. The small subunit is then translocated where it meets the large ribosomal subunit, other protein initiator factors, and initiator tRNA. The tRNA is bound to methionine at a site in the ribosome known as the P site.



Elongation

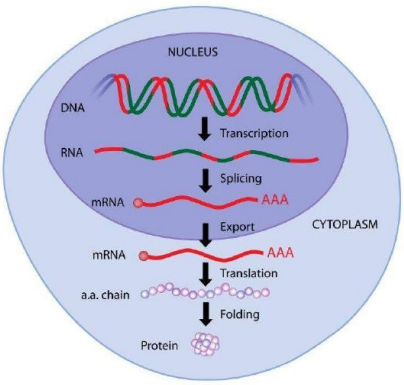
In the elongation phase of protein synthesis, a specific aminoacyl-tRNA, directed by hydrogen bonding interactions between the anticodon region of the aminoacyl-tRNA and the codon region of mRNA, adds to a site distinct from the P site, the A site. The A and P sites are in close proximity, allowing the peptide bond formation between the amino acids. The newly synthesized tRNA bound dipeptide then moves from the A site to the P site. After a translocation of the ribosome in the 5' - 3' direction along the mRNA occurs to expose a new codon. Then, another amino acid-tRNA identity and binds to the mRNA at the A site and the peptidyl transferase reaction is again initiated. As the polypeptide chain grows through subsequent cycles of amino acid residue incorporation, it emerges from the ribosome and undergoes folding into its native secondary and tertiary conformations.

Termination

In termination step, the peptide bond synthesis ceases when a stop codon on the mRNA is reached. This termination site will not bind aminoacyl-tRNA and peptide synthesis stops. Release factors allow the newly synthesized protein to dissociate from the ribosome.

3.5 Gene Expression

Gene expression is a two-step process in which DNA is converted into a protein.



Step 1: The first step is DNA transcription to RNA. In this step, the information from the archival copy of DNA is imprinted into mRNA. The structure of RNA is a little different, it contains ribose instead of deoxyribose, and the four bases that bind to it are cytosine (C), guanine (G), adenine (A), and uracil (U). During transcription, DNA unfolds, and mRNA is created by pairing mRNA bases with the bases of DNA. In this process C in DNA translates to G, G to C, **A to U**, and T to A. After mRNA is transcripted it is transported to the ribosome.

Step 2: The second step, protein translation occurs at the ribosome. During translation, the sequence of codons (triplet of bases) of mRNA is, with the help of tRNA, translated into a sequence of amino acids

Figure 2.6.2.3.5: Gene expression seems to be a straightforward process, the mechanism that control the gene expression that causes most phenotypic differences in organisms.

3.6 Mutation



Mutations are changes in the genetic sequence (DNA or RNA sequence) , and they are a main cause of diversity among organisms. Although some of mutations are beneficial, offering resistance to disease or improved structure and/or function, some other specific mutations can lead to disease and/or death of the cell or organism.

Mutations can occur due to assaults from the environment or spontaneous mutation may occur during the DNA replication. Mutations are estimated to occur at an approximate rate of 1000–1,000,000 per cell per day in the human genome, and every new cell is believed to contain approximately 120 new mutations.

TYPES OF MUTATION

Point mutations when only a single base pair is changed into another base pair. They can be classified as the following:

- **Transition:** When a purine nucleotide is changed to a different purine (A ↔ G) or a pyrimidine nucleotide is changed to a different pyrimidine nucleotide [C ↔ T(U)].
- **Transversion:** When the orientation of a single purine and pyrimidine nucleotide is reversed [A/G ↔ C/T(U)].
- **Silent:** When the same AA is coded.

- **Missense:** When a different AA is coded.
- **Neutral:** When an AA change occurs but does not affect the protein's structure or function.
- **Nonsense:** When a stop codon results, terminating translation and shortening the resulting protein.

Insertion and ***deletion*** mutations, which are together known as ***indels.*** Indels can have a wide variety of lengths. At the short end of the spectrum, indels of one or two base pairs within coding sequences have the greatest effect, because they will inevitably cause a frameshift, i.e. change the entire reading of the mRNA sequence. At the intermediate level, indels can affect parts of a gene or whole groups of genes. At the largest level, whole chromosomes or even whole copies of the genome can be affected by insertions or deletions. At this high level, it is also possible to invert or translocate entire sections of a chromosome, and chromosomes can even fuse or break apart.

If a large number of genes are lost as a result of one of these processes, then the consequences are usually very harmful.

3.7 DNA Repair

The human body have mechanisms to detect and repair the various types of damage that can occur to DNA, no matter whether this damage is caused by the environment or by errors in replication.

Because DNA is a molecule that plays an active and critical role in cell division, during the cell cycle, checkpoint mechanisms ensure that the DNA is intact before permitting DNA replication and cell division to occur. Failures in these checkpoints can lead to an accumulation of damage, which in turn leads to mutations.

UV radiation causes DNA lesions that may distort DNA's structure, introducing bends or kinks and thereby impeding transcription and replication. These lesions may be repaired through a process known as **nucleotide excision repair (NER)**, a mechanism where an enzyme catalyze the removal of damaged nucleotides, and replacement of the correct sequence, guided by the intact complementary DNA strand. Defects in this mechanism is related to human diseases like skin cancer.

Another repair mechanism that handles the spontaneous DNA damage caused oxidation or hydroxylation generated by metabolism is the **base excision repair (BER)**. In this mechanism, enzymes known as DNA glycosylases remove damaged bases by literally cutting them out of the DNA strand through cleavage of the covalent bonds between the bases and the sugar-phosphate backbone. The resulting gap is then filled by a specialized repair polymerase and sealed by ligase.

DNA damage also may occur in form of double-strand breaks, which are caused by ionizing radiation, including gamma rays and X-rays. Double-strand breaks may be repaired through one of two mechanisms: **nonhomologous end joining (NHEJ)**, where an enzyme called DNA ligase IV uses overhanging pieces of DNA adjacent to the break to join and fill in the ends; or **homologous recombination repair (HRR)** where the homologous chromosome itself is used as a template for repair.

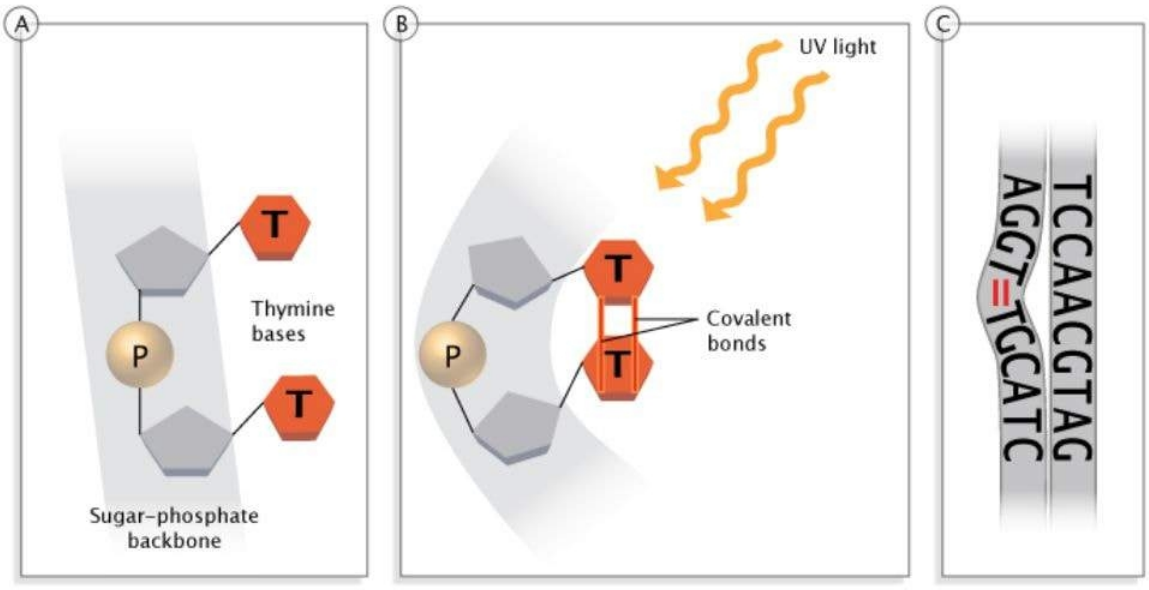


Figure 2.6.2.3.6: UV Radiation can distort DNA, impeding transcription and replication. A region of DNA is shown before (A) and after (B) its

structure is distorted by UV radiation. The distortion in panel B is due to a convex bend in one strand (C). - Source: © Nature Education

Topic 3: Key Points

In this section, we explored the following main points:

1. The process of DNA replication is controlled by several proteins that act together to assure the correct base pairing for creation of the new DNA strand.
2. Transcription is the first step of gene expression, in which a particular segment of DNA is copied into RNA (mRNA) by the enzyme RNA polymerase.
3. In translation, messenger RNA (mRNA)—produced by transcription from DNA—is decoded by a ribosome and tRNA to produce a specific amino acid chain, or protein.
4. Mutations are changes in the genetic sequence (DNA or RNA sequence), that can be beneficial or may result in damage, if not repaired.

Knowledge Check

1. RNA splicing is the process which involves the removal of introns present in the pre-mRNA and splicing of the remaining exons. True or False?

True

False

Answer

true

2. Which of the following does not belong to the process of DNA replication?

Primase

DNA Polymerase

RNA Polymerase

Answer

RNA Polymerase

3. During protein translation, the sequence of codons (triplets of bases) of mRNA is important to:

Maintain the structure of the mRNA

Translated the correct a sequence of amino acids.

DNA replication

Answer

Translated the correct a sequence of amino acids.

This page titled 2.6.2.3: DNA and RNA Metabolism is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **2.3: DNA and RNA Metabolism** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

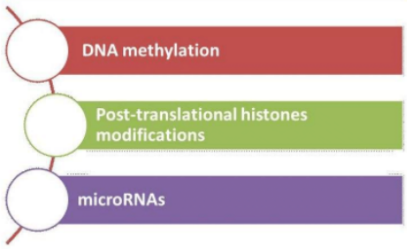
2.6.2.4: Epigenetic Mechanisms

Learning Objectives

- 1: Identify the main epigenetic mechanisms related to control of gene expression.

4.1

Introduction

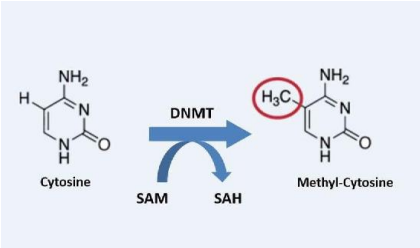


Epigenetics is defined as potentially heritable and reversible changes in gene expression mediated by methylation of DNA, modifications of histone proteins or by non-coding RNAs that are not due to any alteration in the DNA sequence. These processes singularly or jointly affect transcript stability, DNA folding, nucleosome positioning, chromatin compaction, and ultimately nuclear organization. They determine whether a gene is silenced or activated and when and where this occurs.

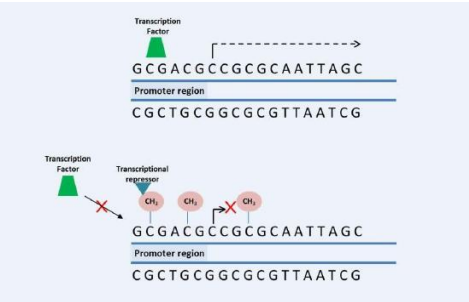
Epigenetic change is a regular and natural occurrence, essential for normal cell development, but can also be influenced by several factors including age, the environment/lifestyle, and disease state.

4.2

DNA Methylation



DNA methylation is a covalent modification of DNA, in which a methyl group is transferred from S-adenosylmethionine (SAM), that is converted to S-adenosylhomocisteine (SAH), to the 5 position of cytosine by a family of enzymes known as DNA methyltransferases (DNMT).Figure 2.6.2.4.1: Disaccharide structure



DNA methylation occurs predominantly in cytosines located in proximity of guanines, known as CpG dinucleotides (CpGs) or CpG island. These CpG island are found in promoter region of genes. The methyl group inhibits the binding of transcription factors to their recognition site, resulting in inhibition of gene transcription. Furthermore, the methyl group may attract other proteins know as transcriptional repressors that contributes for inhibition of gene transcription.Figure 2.6.2.4.2: DNA methylation in the promoter region – Credit Aline de Conti

4.3

Histones Post-Translational Modifications

The nucleosome is composed of five histone proteins (H1, H2A, B, H3, and H4). The N-terminus of these histone proteins are subject to covalent modifications such as methylation, phosphorylation, acetylation, ubiquitination or sumoylation by a group of histone-modifying enzymes . Alterations in these proteins contribute to the accessibility and compactness of the chromatin, and result in activation or suppression of particular genes.

EXAMPLES OF TYPES AND ROLES OF HISTONE MODIFICATIONS

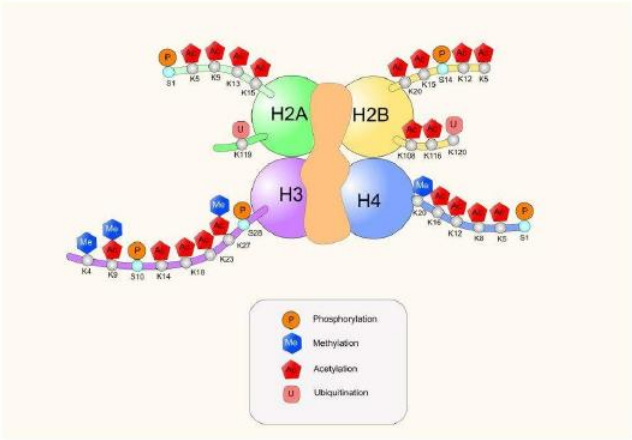


Figure 2.6.2.4.3: Histone sites of post-translational modifications.

Activation of gene Transcription:

Acelylation in the lysine 9(K9) of histone 3(H3) - H3K9ac

Acetylation in the lysine 27(K27) of histone 3(H3) - H3K27ac

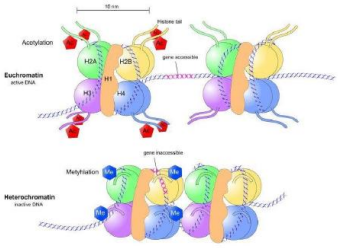
Trimethylation in the lysine 4(K4) of histone 3(H3) - H3K4me3

Repression of gene transcription:

Trimethylation in the lysine 4(K3) of histone 3(H3) - H3K9me3

Trimethylation in the lysine 27(K27) of histone 3(H3) - H3K27me3

Trimethylation in the lysine 20(K20) of histone 3(H3) - H3K20me3



DNA methylation and histone post-translational modifications play important role in the establishment of chromatin structure and in consequence in the gene expression modulation. These epigenetic mechanisms can be independent or can happen together to control the gene expression. They also may influence each other, for example the histone methylation can help to direct DNA methylation patterns, and DNA methylation might serve as a template for some histone modifications after DNA replication.Figure 2.6.2.4.4: Histone acetylation and methylation in gene regulation.

4.4

microRNAs

miRNAs have an important role in gene regulation and they can influence biological functions including cell differentiation and proliferation during normal development and pathological responses. They are small non-coding RNA molecules (containing about 22 nucleotides), derived from regions of RNA transcripts that fold back on themselves to form short hairpins. They regulate gene expression at the post-transcriptional level. A number of miRNAs may bind to specific regions of the messenger RNA (mRNA) and block its translation to proteins. Alteration of the expression of miRNAs is believed to contribute to the progression of tumorigenesis and other diseases.

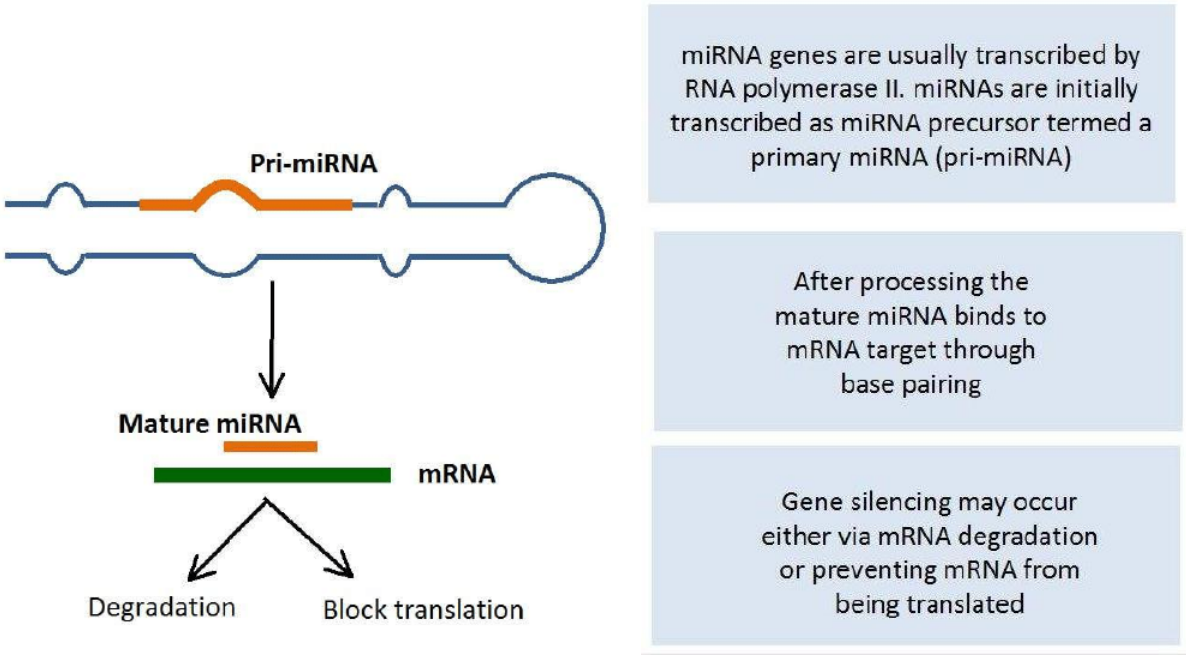


Figure 2.6.2.4.5: Schematic representation of gene silencing by microRNA. Credit: Aline de Conti

Topic 4: Key Points

In this section, we explored the following main points:

- 1: Epigenetic mechanisms may influence gene expression without alteration in the DNA sequence.
- 2: Main epigenetic mechanisms are DNA methylation, histones post-translational modification and alteration in the expression of microRNAs.
- 3: Epigenetic change is a regular and natural occurrence but can also be influenced by several factors including age, the environment/lifestyle, and disease state.
- 4: Different from genetic alterations, epigenetic alterations are considered reversible.

Knowledge Check

1. Epigenetic mechanisms are considered:

Reversible

Irreversible

Answer

Reversible

2. DNA methylation and histone post-translational modifications play important role in the establishment of chromatin structure and in consequence in the gene expression modulation. *True or False?*

False

True

Answer

True

3. DNA methylation may inhibit the binding of transcription factors to their recognition site, at promoter regions resulting in:

Inhibition of gene transcription

Activation of gene transcription

Answer

Inhibition of gene transcription

This page titled 2.6.2.4: Epigenetic Mechanisms is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **2.4: Epigenetic Mechanisms** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.2.5: Section 2 Final Evaluation

1. The carbohydrate component of a nucleoside and nucleotide is usually:

The sugar ribose for RNA molecule and deoxyribose for DNA molecule

Carboxylic group for DNA molecule and amino group for RNA molecule

The deoxyribose for RNA molecule and sugar ribose for DNA molecule

The deoxyribose for RNA molecule and deoxyribose for DNA molecule

Answer

The sugar ribose for RNA molecule and deoxyribose for DNA molecule

2. During protein translation, the sequence of codons (triplets of bases) of mRNA is important for:

Maintaining the structure of the mRNA

Translating the correct sequence of an amino acid

DNA replication

DNA synthesis

Answer

Translating the correct sequence of an amino acid

3. In DNA replication, the lagging strand short chains of nucleic acids, called Okazaki fragments, are generated and the enzyme DNA _____ joins the Okazaki fragments together with lagging strand for replication to proceeds.

topoisomerase

polymerase

helicase

ligase

Answer

ligase

4. Point mutation when the orientation of a single base pair purine and pyrimidine nucleotide is reversed is termed:

Transition

Missense

Transversion

Nonsense

Answer

Transversion

5. The main function of _____ is to control the fluidity of biological membranes.

Glycerol

Cholesterol

Glucose

Phospholipid

Answer

Cholesterol

6. The smooth endoplasmic reticulum (ER) is involved in the synthesis of fatty acids and

Detoxification by hydroxylation reactions

Provides mechanical support for cells

Monooxygenase reaction to produce hydrogen peroxide (H2O2)

Production of energy

Answer

Detoxification by hydroxylation reactions

7. The only organelles which are equipped with their own (circular) DNA, RNA and ribosomes and can perform their own protein synthesis are:

Nucleus

Golgi apparatus

Peroxisomes

Mitochondria

Answer

Mitochondria

8. Phospholipids are the major component of cell membranes forming lipid bilayers because of their amphiphilic property whose structure consist of:

Two hydrophilic fatty acid “tails” and a hydrophobic head consisting of a phosphate group (PO4-3) attached to the third glycerol carbon.

Two hydrophobic fatty acid “tails” and a hydrophobic head consisting of a phosphate group (PO4-3) attached to the third glycerol carbon.

Two hydrophobic fatty acid “tails” and a hydrophilic head consisting of a phosphate group (PO4-3) attached to the third glycerol carbon.

Four hydrophilic fatty acid “tails” and a hydrophobic head consisting of a phosphate group (PO4-3) attached to the third glycerol carbon.

Answer

Two hydrophobic fatty acid “tails” and a hydrophilic head consisting of a phosphate group (PO4-3) attached to the third glycerol carbon.

9. The most important structural component of a nucleus consisting of positively charged protein and negatively charged DNA is called:

Lysosome

Chromatin

Cytoskeleton

Nucleolus

Answer

chromatin

10. The difference between amino acids and fatty acids in terms of structure is that amino acids carry a carboxyl and an amino group while fatty acids carry a carboxyl and a methyl group.

True

False

Answer

true

11. An example of a nucleotide with a two ring structures is:

Cytosine (C)

Adenine (A)

Uracil (U)

Thymine (T)

Answer

Adenine (A) [and also Guanine(G)]

12. DNA methylation occurs predominantly in _____ to inhibit the binding of transcription factors to their recognition site which results in the inhibition of gene transcription.

Guanine

Cytosine

Adenine

Thymine

Answer

Cytosine (C)

This page titled 2.6.2.5: Section 2 Final Evaluation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **Section 2 Final Evaluation** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.3: Principles of Genetic Toxicology

The genetic toxicology methodology or assay technique helps to test or evaluate the level of damages of the genetic information caused by toxicants or agents within the cells. In this e-module you will learn about different genetic toxicology assays, different genetic damages, and cytotoxicity and epigenetics assays.

2.6.3.1: Introduction to Genetic-toxicology Assay

2.6.3.2: Different Genetic Damages or Mutations

2.6.3.3: Different Genetic-Toxicology Assays

2.6.3.4: Different Cytotoxicity Assays

2.6.3.5: Epigenetics Assay

2.6.3.6: Section 3 Final Evaluation

This page titled 2.6.3: Principles of Genetic Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.3.1: Introduction to Genetic-toxicology Assay

Learning Objectives

- Know about the definition of genetic toxicology assay.

1.1. What is Genetic-Toxicology Assay?

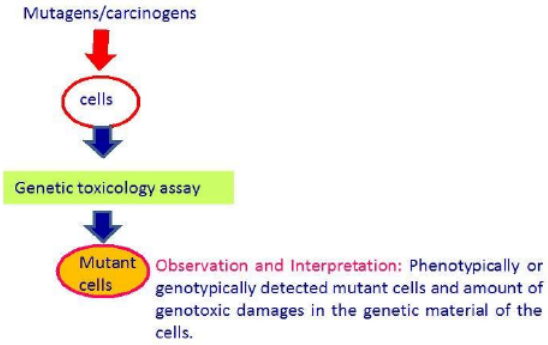


Figure 2.6.3.1.1: Diagrammatic steps in genetic toxicology assay.

The genetic-toxicology methodology or assay technique helps to test or evaluate the level of damage to the genetic information caused by toxicants or agents within the cells (Figure 1). Damages resulted as induced **mutations**, which may lead to different diseases including cancer. The causative toxic agents are know as **mutagens**. Mutagens caused for cancer disease, is known as **carcinogens.**

Topic 1: Key Points

In this section, we explored the following main points:

- 1: Definition of genetic toxicology methodology
- 2: Causative toxic agents as mutagens.
- 3: Genotoxic damages result in the phenotypically or genotypically mutant cells.

Knowledge Check

Genetic-toxicology methodology or assay technique helps to test...

the level of damages caused by toxicants within the cells.

the level of reactive oxygen species (ROS) within the cells.

the stages of cancer disease.

Answer

the level of damages caused by toxicants within the cells.

This page titled 2.6.3.1: Introduction to Genetic-toxicology Assay is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1: Introduction to Genetic-toxicology Assay** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.3.2: Different Genetic Damages or Mutations

Learning Objectives

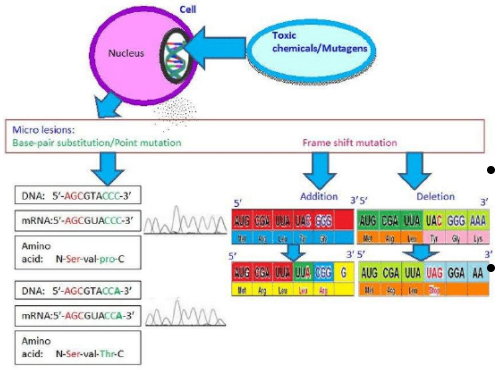
- 1: Know different types of genetic damages or mutations.
- 2: Know how different mutations resulted in the cells by mutagens.

Do you know how many mutations or damages are possible by mutagens in the cells?

Mutations can be:

- Microlesions (gene mutation)
- Macro lesions (Chromosomal mutation)

2.1: Microlesions (Gene Mutation)



Microlesions are the damages or mutations in DNA bases. These mutations are with invisible phenotypic changes. The types of microlesions are illustrated in the figure on the right.

Base-pair substitution mutation (qualitative change in nucleotide pairs)

- In this mutation, single base nucleotide is replaced by another nucleotide.

Frame shift (quantitative change in nucleotide pairs)

- In this mutation, addition or deletion of nucleotide in the DNA sequence resulted to shift or change the entire DNA or amino acid sequence.

Figure 2.6.3.2.1: Micro lesions mutation by mutagens.

2.2 Macro Lesions (Chromosomal Mutation)

Macro lesions are chromosomal mutations with mutagens and are with distinct morphological changes in the phenotype. These morphological changes of chromosomes can be cytologically visible under microscope. Macro lesions are following types:

Numerical changes in chromosomes

- Polyploidy: Duplication of entire set of chromosome to triploid or tetraploid.
- Aneuploidy: Changes of single missing chromosome to **monosomy** or three copies of a single chromosome to **trisomy**.

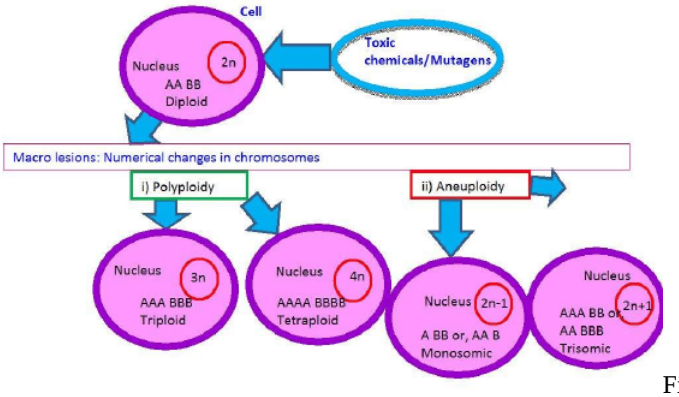


Figure 2.6.3.2.2: Macro lesions mutation by mutagens.

Structural changes in chromosomes

- Deletion: loss of chromosome segment

- Translocation: A segment of one chromosome becomes attached to a non homologous chromosome. It can be one way transfer as simple translocation and two way transfer as reciprocal translocation.
- Inversion: A change in the direction of material along a single chromosome.
- Duplication: Repetition of chromosome segment

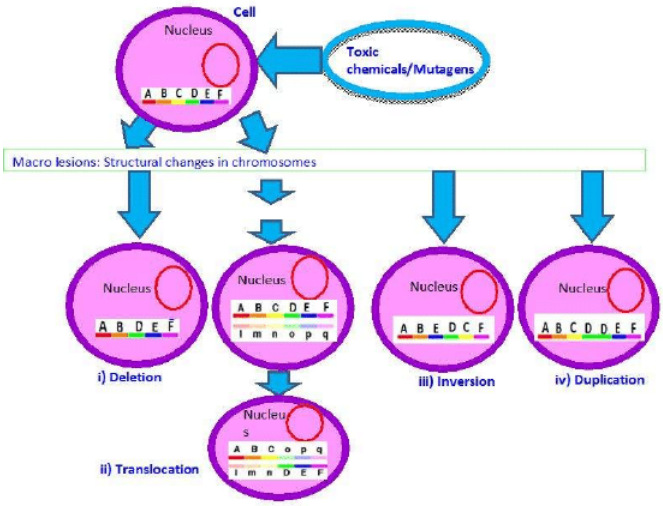


Figure 2.6.3.2.3: Structural changes in chromosomes resulted to macro lesions mutation.

Micronuclei changes

- Micronuclei (MN) are the damaged chromosome fragments or whole chromosomes that were not incorporated into the cell nucleus and stayed as the extra-nuclear bodies after the cell division.
- MN can be resulted by the defects of the cell repair machinery and by the accumulation of damaged DNA and chromosomal aberrations.genetic.html

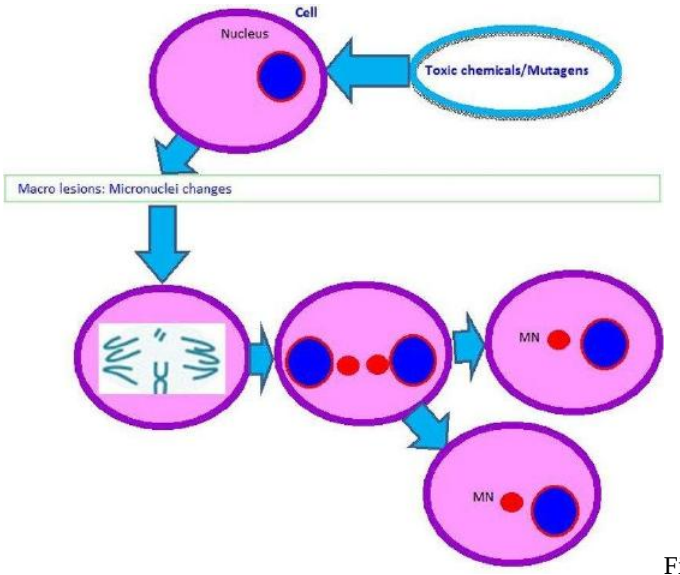


Figure 2.6.3.2.4: Micronuclei changes resulted to macro lesions mutation.

Topic 2: Key Points

In this section, we explored the following main points:

- 1: Different types of Microlesions (gene mutation) and Macro-lesions (Chromosomal mutation).

- 2: How toxic agents or mutagens modulate the two different microlesions namely base pair substitution and frame shift mutation.

- 3: Genotoxic mutagens involve also in different macro-lesions namely numerical or structural changes in chromosomes or changes of micronuclei in cell.

Knowledge Check

1. What is monosomy?

Duplication of entire set of chromosome .

Single missing chromosome from diploid set.

Three copies of a single chromosome.

Answer

Single missing chromosome from diploid set.

2. In base-pair substitution mutation, single base nucleotide is replaced by another nucleotide.

True

False

Answer

True

3. What are the structural changes in chromosomes caused by toxicants?

Deletion

Translocation

Inversion

Duplication

All of the above

Answer

All of the above

4. Micronuclei (MN) changes are the damaged chromosome fragments or whole chromosomes that were not incorporated into the cell nucleus and stayed as the extra-nuclear bodies after the cell division.

True

False

Answer

true

5. Frame shift mutation, resulted to shift or change of entire DNA or amino acid sequence by:

addition of nucleotide in the DNA.

deletion of nucleotide in the DNA.

addition or deletion of nucleotide in the DNA

Answer

addition or deletion of nucleotide in the DNA

This page titled 2.6.3.2: Different Genetic Damages or Mutations is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **3.2: Different Genetic Damages or Mutations** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.3.3: Different Genetic-Toxicology Assays

Learning Objectives

- 1: Know different types of genetic-toxicology assays
- 2: Know how different genetic-toxicology assays are used in toxicology when cells are exposed to mutagens.

The goal of genetic toxicology assay is to determine whether any chemical or mutagen will do any adverse effect on genetic material or may cause different diseases including cancer. The assays can be performed using bacterial, yeast, or mammalian cells. One can early control and save vulnerable organisms from genotoxic chemicals by preforming genetic toxicology assay.

The following different types of genetic toxicology assays are used now a days:

- Bacterial Reverse Mutation Assay (Ames Assay)
- Genetic mutation assay
  - Allele-Specific PCR
  - Sanger Dideoxy Sequencing
- Chromosome aberration study
- Micronucleus assay

3.1: Bacterial Reverse Mutation Assay (Ames Assay)

This assay was discovered by Bruce Ames in 1970. This assay is widely used to test for gene mutation. The technique uses several strains of the bacterium Salmonella typhimurium which carry mutations in genes involved in histidine synthesis. These strains are auxotrophic mutants and they require histidine for growth and they cannot produce it. This assay examine the ability of the chemical or mutagen in creating mutations or a "prototrophic" state of strains, when the strains can grow on a histidine-free medium.

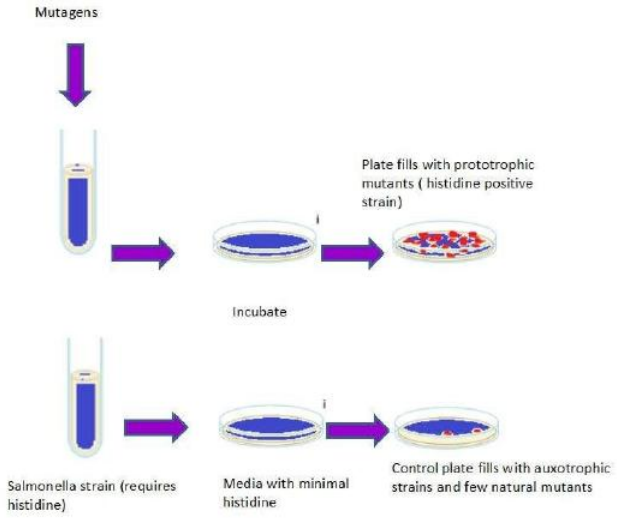


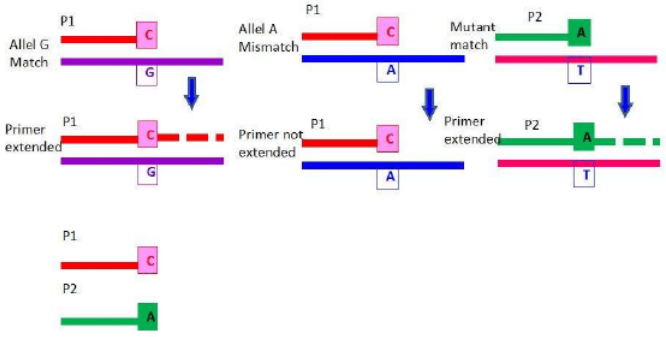
Figure 2.6.3.3.1: Ames test

3.2: Genetic Mutation Assay

Following are the different molecular assay to study nucleotide variants or alternation of genetic material caused by mutagens:

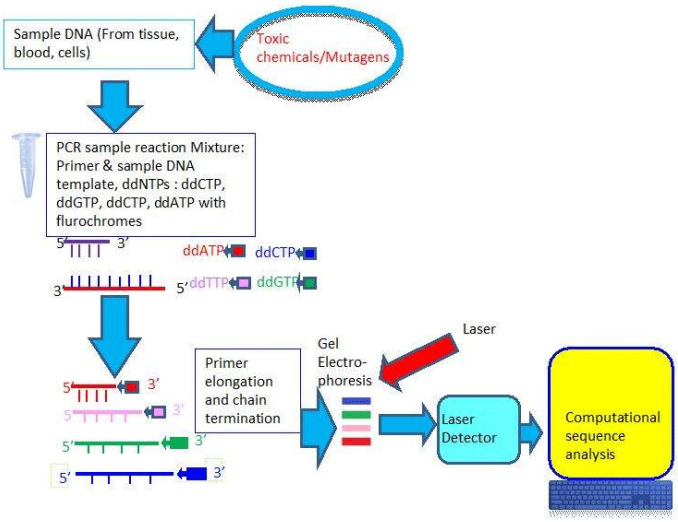
Allele-Specific PCR

Single nucleotide polymorphism (SNP) resulted from base substitution mutation can be analyzed by this method. In this real-time PCR, fluorescent reporter probes are added to the reaction mixture and one fluorescent reporter probe is selected for wild type and other fluorescent probe is used for mutant. The PCR primers with fluorescent probe will match or mismatch one of the alleles at 3' end of the primer. DNA polymerase extends the probes in a complementary fashion and releasing the reporter fluorescent molecules for detection. The PCR cycles with the reporter probes show the amplified signals and allow for precise measurement of one or both alleles of interest. Similarly, the 3' end of the mutant-specific primer is extended only in the presence of DNA with that mutation.



Sanger Dideoxy Sequencing

The goal of this method is to detect unknown mutations including single nucleotide variants (SNVs) and small duplications, insertions, deletions, and indels of interest caused by mutagens. In this method, sequencing primers hybridized to the PCR product and are extended using the four deoxynucleotides (dNTPs), a mixture of fluorescently labeled dideoxynucleotides (ddNTPs) and DNA polymerase. Four ddNTPs are marked with a different fluorescent dye. Random incorporation of the marked ddNTPs shows in termination of strands at each location along the sequence. The gel electrophoresis separates the strands by size. Fluorescence spectroscopy measured the terminating nucleotides.



3.3: Chromosome Aberration Study

Cytogenetic assays of mammalian cells are performed to detect different types of structural and numerical chromosomal aberrations caused by genotoxic chemicals. The clastogenic or aneugenic effects from the genotoxic chemicals will result in an increase in frequency of structural (premature centric separation, chromosome breaks, dicentric chromosomes, ring) complex rearrangements (Figure 4) or numerical aberrations of the genetic material in mammalian cells.

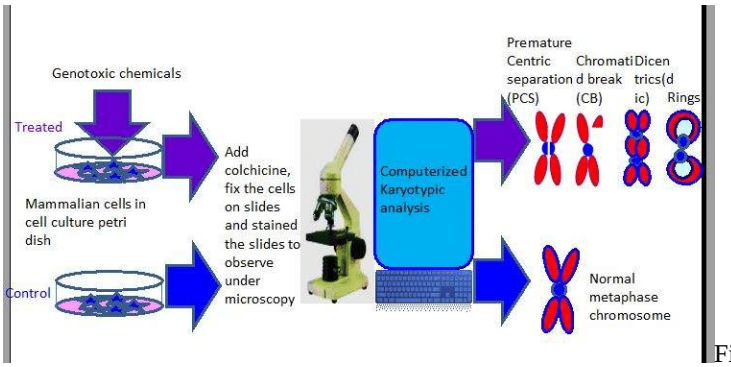


Figure 2.6.3.3.2: Chromosome aberration study

3.4: Micronucleus Assay

Micronucleus assay is used as a tool to evaluate genetic damage caused by genotoxic chemicals. The number of micronuclei (Figure 5) generated directly relates to the amount of DNA damage in the cells.

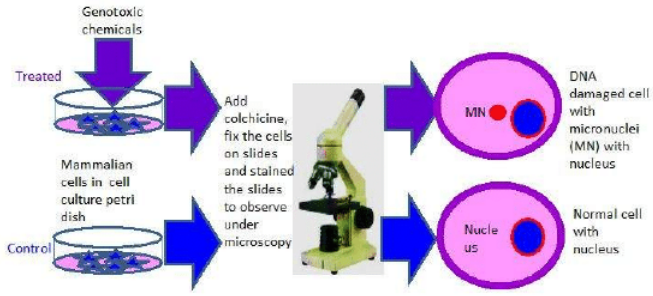


Figure 2.6.3.3.3: Micronucleus assay

Topic 3: Key Points

In this section, we explored the following main points:

- 1: Definition of genetic toxicology assay
- 2: Different types of genetic-toxicology assays.
- 3: How genetic mutation assays are performed by Allele-Specific PCR and Sanger Dideoxy Sequencing techniques.
- 4: What are the different types of chromosomal aberrations observed under microscope by chromosome aberration study?
- 5: Changes of micronuclei observed under microscope by Micronucleus assay.

Knowledge Check

1. Cytogenetic assays of mammalian cells are performed to detect different types of structural and numerical chromosomal aberrations caused by a genotoxic chemicals. The structural chromosomal aberrations are:

premature centric separation

ring

chromosome breaks

dicentric chromosomes

All of the above

Answer

All of the above

2. In Allele-Specific PCR, fluorescent reporter probes are added to the reaction mixture and one fluorescent reporter probe is selected for wild type and other fluorescent probe is used for mutant.

True

False

Answer

True

3. Which instrument is used to measure the terminating nucleotides in Sanger Dideoxy Sequencing?

Fluorescence spectroscopy

Spectrophotometer

Fluorescence microscopy

None of the above

Answer

Fluorescence spectroscopy

4. The Ames technique uses several strains of the bacterium Salmonella typhimurium which carry mutations in genes involved in:

Arginine Synthesis.

Histidine synthesis.

Lysine Synthesis.

None of the above.

Answer

Histidine synthesis.

This page titled 2.6.3.3: Different Genetic-Toxicology Assays is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **3.3: Different Genetic-Toxicology Assays** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.3.4: Different Cytotoxicity Assays

Learning Objectives

- 1: Know different types of cytotoxicity assays.
- 2: Know how different cytotoxicity assays are used when cells are exposed to toxicants or mutagens.

The goal of cytotoxicity assay is to determine whether any chemicals or drugs will do any toxic effect or load on milieu or genetic material of the cells caused for lethality of the cells or caused for different diseases.

The following are the different types of cytotoxicity assays:

- DNA fragmentation/ladder assay
- Comet assay
- Necrosis assay
- Enzyme assay
- Proteomics assay
- Expression array assay

4.1: DNA Fragmentation/Ladder Assay

DNA fragmentation or ladder assay are used to know the fragmented DNA of the cells caused by chemicals or drugs. Fragmented DNA can be separated by agarose gel electrophoresis and can be visualized as “ladder” by ethidium bromide staining. The evaluation of cytotoxicity through cell death is an acceptable common assessment.

Ladder assays are performed for the following reasons:

1. to simply characterize the toxicity of the chemicals or drugs in cells, or
2. to determine the maximum doses of the test chemicals or drugs that can be used for cells without causing too much cell death.

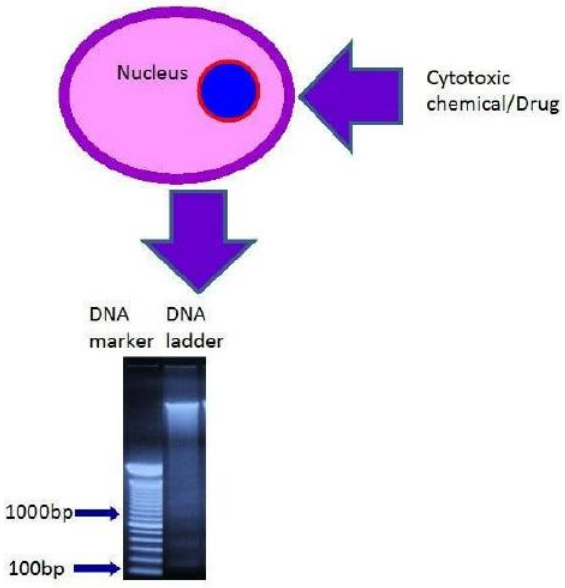


Figure 2.6.3.4.1: DNA ladder. Nucleotide base pair mentioned as,“bp”.

4.2: Comet Assay

The Comet Assay (single cell gel electrophoresis /SCGE) is used to detect DNA damage by using a micro gel electrophoresis. The image of the damaged DNA shows a comet with head and tail. The analysis of image for comet assay is calculated for the “tail length” of the comet which is the measurement from the point of highest intensity within the comet head as well as the “tail moment” which is the product of the tail length and the fraction of total DNA present within the tail.

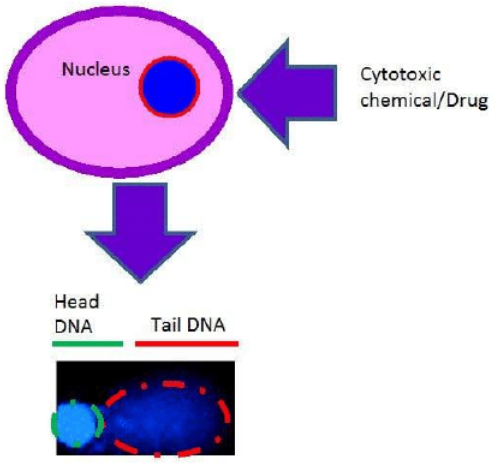


Figure 2.6.3.4.2: Comet assay showed amount of damaged DNA by comet tail moment.

4.3: Necrosis Assay

The necrosis assay is performed by flow cytometry analysis with staining of Annexin V and propidium iodide (PI) in the cells. The cells are considered in the stage of necrosis if the cells lose membrane integrity and die promptly due to cell lysis when exposed to chemicals, drugs, toxins or foreign antigens. In necrosis, cells show swelling, loss of membrane integrity and disruption of metabolism. Cells with necrosis do not go to stage of apoptosis, apoptotic cell may undergo secondary necrosis. These necrotic cells will shut down metabolism, lose membrane integrity, lyse and formed cell injury autolysis.

The flow cytometric analysis for necrosis assay showed that the cells stained positive for both FITC Annexin V and PI are in the end stage of apoptosis and are undergoing to the stage of necrosis as dead cells stained PI positive. Cells that stain negative for both FITC Annexin V and PI are alive and not undergoing apoptosis or necrosis.

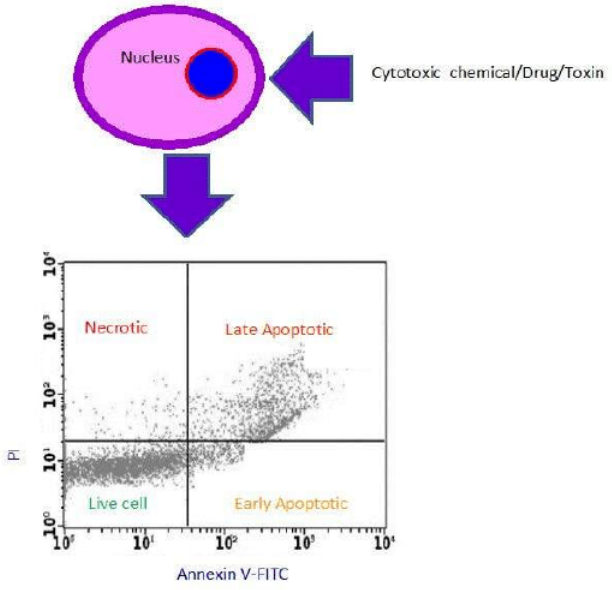


Figure 2.6.3.4.3: Flow cytometry analysis of cells showing necrosis stage.

4.4: Enzyme Assay

The enzyme assay is used to monitor passaging of lactate dehydrogenase (LDH), due to loss of cell membrane integrity when cells are exposed to cytotoxic compounds. LDH reduces NAD to NADH which generates a color change by interaction with a specific

probe (Figure 4a). In other enzyme assay, Adenosine triphosphate (ATP)-based assay combined with bioluminescent assay are used to measure cytotoxicity of the cells in which ATP is the reagent for the luciferase reaction.

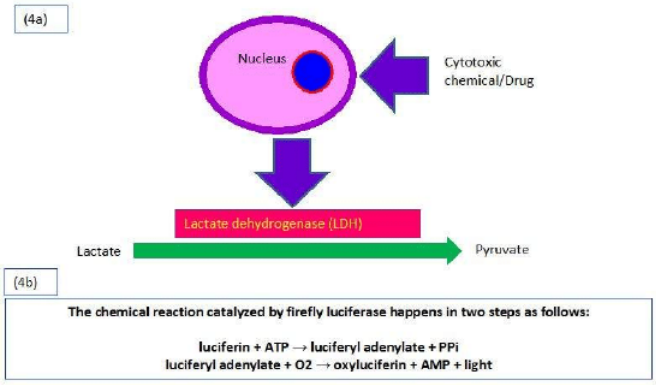


Figure 2.6.3.4.4: Schematic presentation of enzyme assay. (a) LDH assay (b) ATP based assay

4.5: Proteomics Assay

The proteomic assay is performed to know the mechanism of cellular toxicity by measuring expression of a specific protein which may consider as a biomarker for particular toxic mechanism or cellular toxicity signaling pathway. Immunofluorescence, immunoprecipitation and immunoblot assay are mainly used to know the effect of toxicants in cellular toxicity signaling pathway or mechanism.

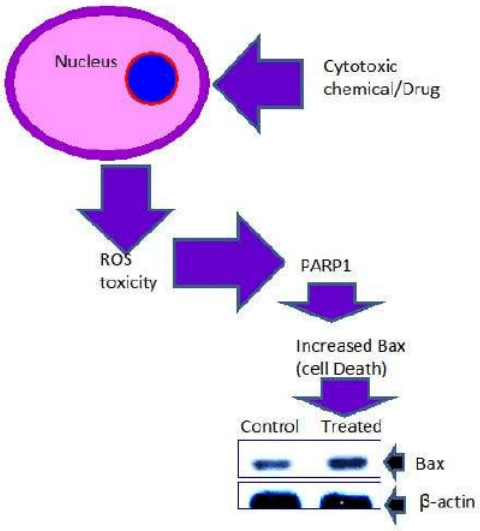


Figure 2.6.3.4.5: Schematic presentation of proteomic assay.

4.6: Expression Array Assay

The expression array is the chip based microarray of more gene expressions (finger print of genes) by the effect of cellular toxicants. This is a rapid and sensitive detection method which allows to detect all toxicological end points at wide range of molecular level changes in the cell at single assay. The microarray process can be divided into two main parts. First is the printing of known gene sequences onto glass slides or other solid support followed by hybridization of fluorescently labeled cDNA (containing the unknown sequences to be interrogated) to the known genes immobilized on the glass slide. After hybridization, arrays are scanned using a fluorescent microarray scanner. Analyzing the relative fluorescent intensity of different genes provides a measure of the differences in gene expression.

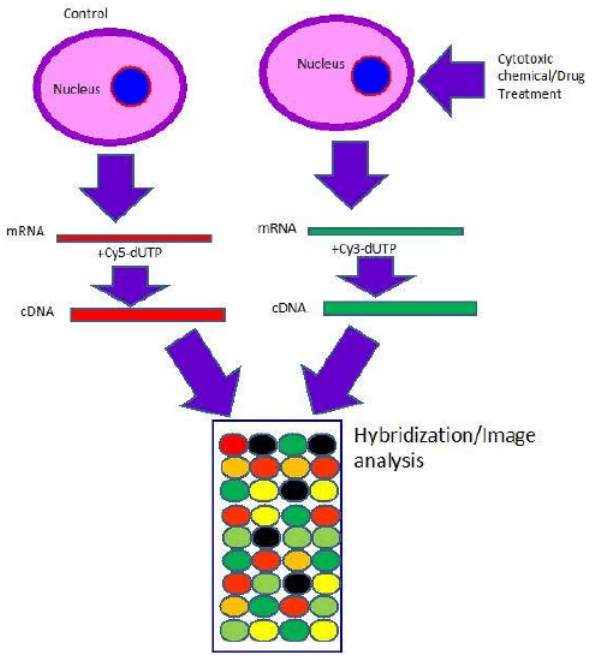


Figure 2.6.3.4.6: Schematic representation of gene expression array showed that Cy5(red) and Cy3 (green) labeled cDNA hybridized to a DNA microarray. Yellow spots indicate the genes are expressed in both samples. The intensity and different types of color at each spot indicate the level and presence of genes in samples. Black spots show low level of expression or do not show any expression of genes.

Topic 4: Key Points

In this section, we explored the following main points:

- 1: Different types of cytotoxicity assays.
- 2: How different cytotoxicity assay namely DNA fragmentation/ladder assay, Comet assay, Necrosis assay, Enzyme assay, Proteomics assay, and Expression array assay are used when cells are exposed to cytotoxic agents.

Knowledge Check

1. The cells are considered in the stage of necrosis, if the cells lose membrane integrity and die promptly due to cell lysis when exposed to chemicals, drugs, toxins or foreign antigens.

True

False

Answer

True

2. The Comet Assay is used to detect DNA damage by using a micro gel electrophoresis:

True

False

Answer

True

3. The expression array is the chip based microarray of more gene expressions (finger print of genes) by the effect of cellular toxicants.

True

False

Answer

True

4. Ladder assay are performed for the following reasons: 1) to simply characterize the toxicity of the chemicals or drugs in cells, or 2) to determine the maximum doses of the test chemicals or drugs that can be used for cells without causing too much cell death.

True

False

Answer

True

5. What are the proteomic assays to know the effect of toxicants in cellular toxicity signaling pathway or mechanism?

Immunofluorescence assay

ring

Immunoprecipitation assay

Immunoblot assay

All of the above

Answer

All of the above

This page titled 2.6.3.4: Different Cytotoxicity Assays is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **3.4: Different Cytotoxicity Assays** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.3.5: Epigenetics Assay

Learning Objectives

- 1: Know different types of epigenetic assays.
- 2: Know how different epigenetic assays are used when cells are exposed to toxicants or mutagens.

5.1: DNA Methylation Assay

DNA methylation assays are important to know the epigenetic modification which is a heritable, enzyme-induced modification without alteration of the nucleotide base pairs. The transfer of a methyl-group to the 5-carbon on the cytosine in a CpG dinucleotide happens in the DNA methylation by DNA methyltransferases (DNMT1, DNMT3A, and DNMT3B). The high level of promoter CpG island methylation results in gene silencing. The methylated DNA immunoprecipitation (MeDIP)-chip technique is used for DNA methylation assay.

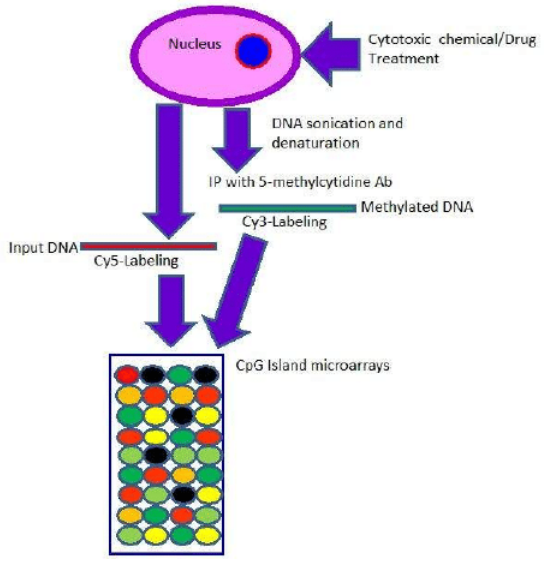


Figure 2.6.3.5.1: Schematic diagram of MeDIP protocol for DNA methylation assay.

In brief, the MeDIP-chip procedure is mentioned as follows. The genomic DNA is sheared to low molecular weight fragments (approximately 400 bp) by sonication. Then, the methylated DNAs are immunoprecipitated with the anti-methyl-cytosine antibody, and are amplified with PCR, if source material is less. Input and methylated DNA are labeled with fluorescent dyes Cy3 (green) and Cy5 (red), pooled, denatured, and are hybridized to a microarray slide containing all the annotated human CpG islands or other whole genome or promoter microarray designs. Then the slide is scanned using a scanner and each image is analyzed with the image analysis software (Figure 1).

5.2: Histone Modification Assay

Histone modification assays are useful to find the modification of histone proteins (e.g. lysine acetylation, lysine and arginine methylation, serine and threonine phosphorylation, and lysine ubiquitination and sumoylation) which have important roles in epigenetic inheritance. The chromatin immunoprecipitation (ChIP) assay followed by hybridization to microarrays (ChIP-chip) (left) or by high-throughput sequencing (ChIP-seq) (right) are both powerful techniques to find histone modification.

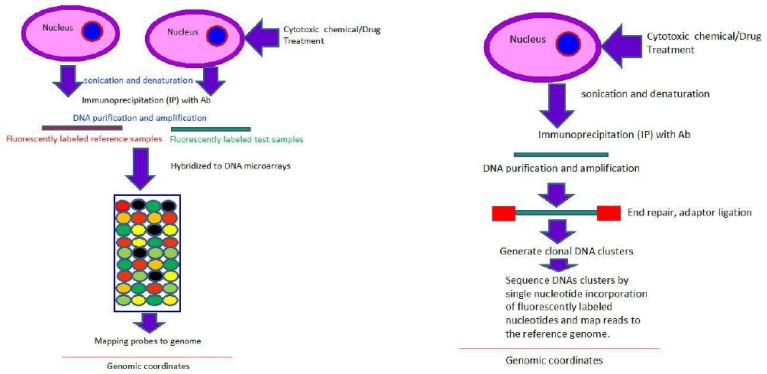


Figure 2.6.3.5.2: Schematic presentation of ChIP – chip microarray and sep of histone modification assay

5.3: MicroRNAs Assay

MicroRNAs assays are used to know the non-coding RNAs (17-25 nucleotides) which target messenger RNAs (mRNAs) and decayed the mRNAs or downregulated at the level of translation into protein. Almost, 60% of human protein coding genes are controlled by miRNAs and these miRNAs are epigenetically regulated. About 50% of miRNA genes are related with CpG islands, which may be repressed by epigenetic methylation. Other miRNAs are epigenetically controlled by either histone modifications or by DNA methylation. The expression of microRNAs are quantified by RT-PCR followed by quantitative PCR (qPCR). Then, miRNAs are hybridized to microarrays, slides or chips with probes to hundreds or thousands of miRNA targets. The microRNAs can be both invented and profiled by sequencing methods (microRNA sequencing).

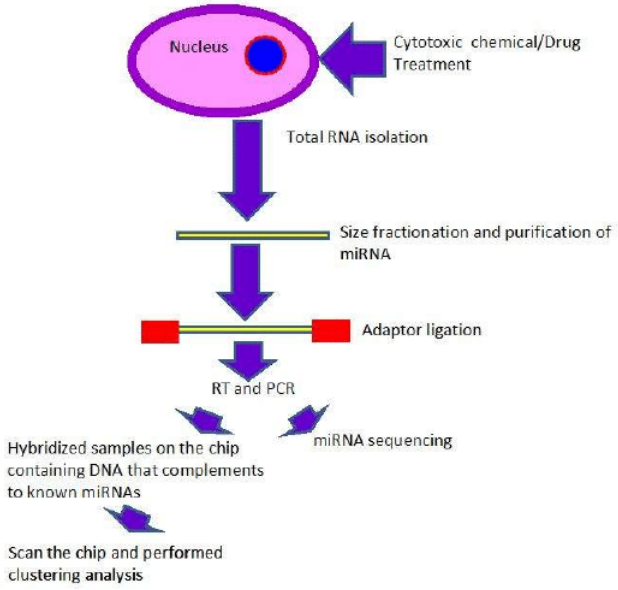


Figure 2.6.3.5.3: Schematic presentation of miRNA assay in epigenetic toxicology

Topic 5: Key Points

In this section, we explored the following main points:

- 1: Different types of Epigenetic assays.

- 2: How different Epigenetic assays namely DNA methylation assay, histone modification assay and MicroRNAs assay are used when cells are exposed to toxic chemicals or agents.

Knowledge Check

1. DNA methylation assays are important to know the non-epigenetic modification.

True

False

Answer

false

2. MicroRNAs assays are used to know the non-coding RNAs. These non-coding RNAs are :

17 to 25 nucleotides.

50 to 100 nucleotides.

200 to 400 nucleotides.

None of the above.

Answer

17 to 25 nucleotides.

3. Histone modification assays are useful to find the modification of histone proteins which have important roles in epigenetic inheritance:

True

False

Answer

true

This page titled 2.6.3.5: Epigenetics Assay is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **3.5: Epigenetics Assay** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.3.6: Section 3 Final Evaluation

1. _____ is a method used in analyzing base substitution mutation resulting from single nucleotide polymorphism.

Micronucleus assay

Chromosome aberration study

Allele-Specific PCR

Chromatography

Answer

Allele-Specific PCR

2. Proteomic assay analyzes the effect of toxicants in cellular toxicity signaling pathways or mechanisms through:

Immunofluorescence

Immunoblot

Immunoprecipitation

All of the above

Answer

All of the above

3. In Ladder Assay, fragmented DNA can be separated by agarose gel electrophoresis and can be visualized as “ladder” by _____ staining.

Ethidium bromide

Eosin

Gram

Wright’s

Answer

Ethidium bromide

4. Trisomy is a form of aneuploidy interpreted as:

Single missing chromosome from diploid set.

Three copies of a single chromosome from a diploid set

Three copies of a single chromosome from a triploid set

Single missing chromosome from a triploid set

Answer

Three copies of a single chromosome from a diploid set

5. Gene mutations in which a single base nucleotide is replaced by another nucleotide are known as:

Frame shift

Quantitative change in nucleotide

Qualitative change in nucleotide

Left shift

Answer

Qualitative change in nucleotide

6. Some structural chromosomal aberrations caused by genotoxic chemicals in cytogenetic assays of mammalian cells include the following EXCEPT:

Dicentric Chromosomes

Ring Chromosomes

Spiral Chromosomes

Chromosome breaks

Answer

Spiral Chromosomes

7. Structural changes in chromosomes include the following EXCEPT:

Aneuploidy

Inversion

Translocation

Deletion

Answer

Aneuploidy

8. Addition or deletion of nucleotides in the DNA sequence results in the change of the entire DNA or amino acid sequence. This process is known as:

Frame shift

Base-pair substitution mutation

Qualitative change in nucleotide

Right shift

Answer

Frame shift

9. In Allele-Specific PCR, fluorescent reporter probes are added to the reaction mixture, one fluorescent reporter probe is selected for the wild type and the other fluorescent probe is used for the mutant.

True

False

Answer

true

10. The analysis of image to determine DNA damage in comet assay is calculated for the _____ and _____.

“tail length” and “head length”

“head length” and “head moment”

“head length” and “tail moment”

“tail length” and “tail moment”

Answer

"tail length" and "tail moment"

11. Cells which stain negative for both fluorescein isothiocyanate Annexin V and propidium iodide in flow cytometric analysis for necrosis assay are:

Dead and undergoing apoptosis and necrosis

Alive and not undergoing apoptosis or necrosis

Dead and undergoing necrosis only

Alive and undergoing apoptosis only

Answer

Alive and not undergoing apoptosis or necrosis

12. Histone Modification Assay uses Chromatin immunoprecipitation assay (ChIP) followed by:

Hybridization to microarrays (ChIP-chip)

Immunofluorescence

Fluorescence spectroscopy

Fluorescence microscope

Answer

Hybridization to microarrays (ChIP-chip)

This page titled 2.6.3.6: Section 3 Final Evaluation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **Section 3 Final Evaluation** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.4: Applied Systems Toxicology

Systems toxicology is a branch of science that utilizes data from different branches of toxicology and integrates them to provide a holistic approach for safety assessment. In this e-module you will learn about the concept of systems toxicology, dose level in toxicology, and different approaches to traditional and new toxicology.

2.6.4.1: Systems Toxicology

2.6.4.2: Dose Level and Applied Toxicology

2.6.4.3: Tools and Technologies in Systems Toxicology

2.6.4.4: Other Approaches for Predictive Toxicity Modeling

2.6.4.5: Technologies Used In Systems Biology/Toxicology

2.6.4.6: Takeaways Summary

2.6.4.7: Section 4 Final Evaluation

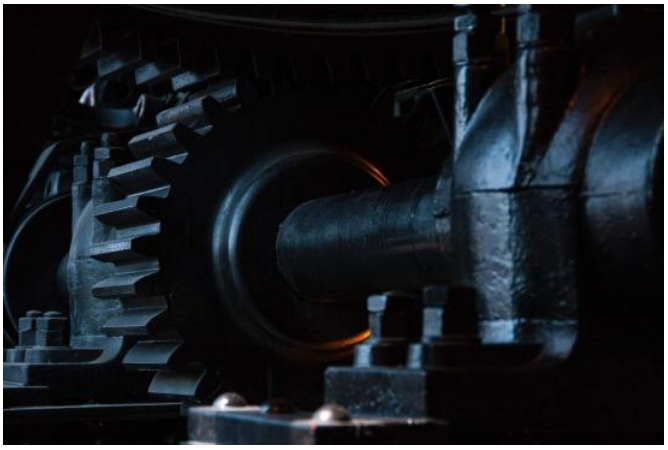
This page titled 2.6.4: Applied Systems Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.4.1: Systems Toxicology

Learning Objectives

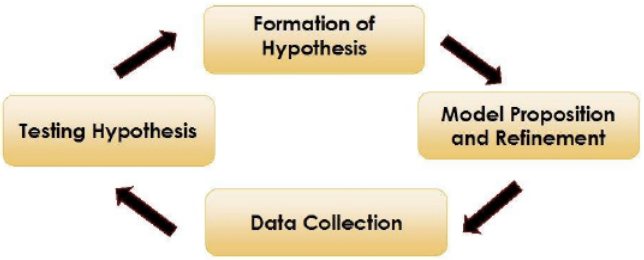
- 1: Understand the concept of systems toxicology.
- 2: Understand the approaches of traditional toxicology approaches vs. the new toxicity testing paradigm.
- 3: Recognize the driving force behind the growth of this field.
- 4: Applications of this field.

The word “systems” originates from the Latin word “systema” which means a complete concept that has several parts. Similarly systems toxicology is a branch of science that utilizes data from different branches of toxicology and integrates them to provide a holistic approach for safety assessment.



Toxicology is the science of understanding the adverse effects of xenobiotics (drugs, chemicals, etc.) on biological systems. Biological systems are extremely complex. Due to the vast number of toxicology research approaches over the years, and a lot of data have been generated in different systems- in vivo, in vitro, in silico (especially due to “omics” approaches). However, there is currently a lack of interpreting/utilizing that data for efficient safety assessment of xenobiotics.

Systems toxicology aims to fill this gap and utilize these data from different systems and integrate them into meaningful assessment for safety. It relies heavily on mathematical and computational models to link the data from various systems. So, in order to have fully validated systems toxicology approaches it is important to have “real” (in life) data from animal models to validate the hypothesis.



The Driving Force

The main driving force behind the development of systems toxicology approaches is the fact that the whole “safety assessment” process is a very lengthy, time consuming and expensive process in case of chemicals as well as the pharmaceutical industry. In order to make this process more efficient, it is important for early pharmaceutical/chemical (especially pharmaceuticals)candidate selection/screening. Screening thousands of compounds is a lengthy process and current high-throughput screening approaches together with large volume data analysis techniques have helped in more efficient selection of target molecules.



Schematic for the drug development process

Topic 1: Key Points

In this section, we explored the following main points:

- 1: The concept of systems toxicology and the different approaches and the driving force behind the development of this field.

Knowledge Check

1. Systems toxicology is usually applicable in the ...

Early discovery phases of drug development

During marketing of the drug

During regulatory safety testing phases

None of the above

Answer

Early discovery phases of drug development

2. Systems toxicology approach involves...

Traditional animal experiments

Alternative in vitro methods

Computational and mathematical models

All of the above

Answer

All of the above

This page titled 2.6.4.1: Systems Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **4.1: Systems Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.4.2: Dose Level and Applied Toxicology

Learning Objectives

After completing this lesson, you will be able to:

- 1: Assess toxicity in response to dose levels.

As students of toxicology it is critically important to understand that everything including water and oxygen has the potential to act as poison. It is only the dose that determines the toxic/beneficial effect.

Dose level and Applied Toxicology

While traditional toxicology approaches in many academic laboratories and also in industry utilized very high dose levels of xenobiotics in order to study various mechanisms of toxicity, it has been identified recently that this is not a very ideal approach especially for industrial applications.

Step 1 Title

- Using very high dose levels may saturate the physiological processes in a living system and hence may cause “forced toxicity” which is not very relevant in terms of real life exposures (could be valid in case of overdose/accidental overexposure).

In industry, a lot of preliminary research is conducted in order to determine dose levels for toxicology experiments. Typically, in case of pharmaceutical compounds a lot of pharmacokinetic modeling and simulations is performed to come up with exposure levels that are multiples of the real drug exposure level in humans. In the chemical industry relevant exposure levels that human beings may be exposed to based on the use of the chemical is determined.

Dose Level Selection

- 1: Traditional toxicity testing involved using a large number of animals and using very high dose levels. Study designs such as LD 50 (Lethal Dose for 50% animals in the study) are no longer used. It is not feasible from a scientific or animal welfare point of view

- 2: Nowadays, dose level selection often involves mathematical simulation and modeling utilizing various forms of in vitro data that are analyzed with the help of medium and high-throughput modeling tools. This is a systems toxicology approach where data from different platforms is utilized to come up with relevant dose levels that are scientifically justifiable and also to refine and minimize animal experiments.

Schematic Approach: Systems Toxicology

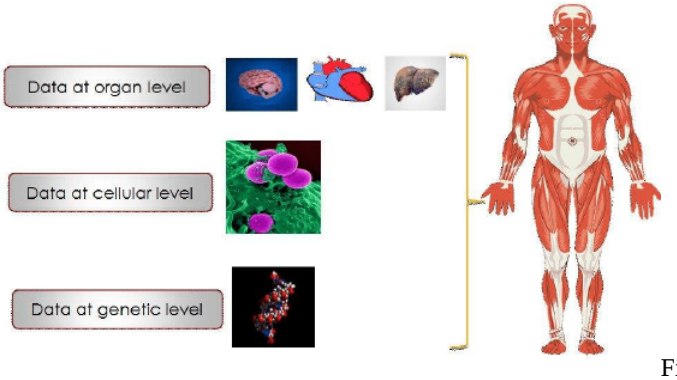


Figure 2.6.4.2.1: Schematic Approach: Systems Toxicology. Credit: pixabay.com

Key Points

In this section, we explored the following main points:

- 1: The understanding of how toxicity is driven strongly by the dose response relationship.

Knowledge Check

1. Utilizing very high dose levels in animal models are not useful because

It causes forced toxicity due to saturation of the absorption and elimination processes

It is too toxic and no meaningful information is available from such data

It is necessary for planning acute studies

It causes a lot of wastage of test article

Answer

It causes forced toxicity due to saturation of the absorption and elimination processes

2. Traditional LD 50 studies evaluated...

Adverse effects in 50 animals per dose group

Death in 50 % of the population

Effects in 50% of the animal population

Doses to be used in subsequent chronic studies

Answer

Death in 50% of the population

This page titled 2.6.4.2: Dose Level and Applied Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **4.2: Dose Level and Applied Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.4.3: Tools and Technologies in Systems Toxicology

Learning Objectives

- 1: Understand tools and technologies used in systems toxicology, the concept of adverse effects and Adverse Outcome Pathways.

Tools and Technologies



1. Tools at the molecular level: Omics technologies such as genomics, protemics, metabolimics.

2. Tools at the cellular level include highly refined high-throughput *in vitro* model systems.

3. Modeling softwares that can utilize in vitro data and mathematically translate that into relevant in vivo information utilizing pharmacokinetic (PK) or physiologically based pharmacokinetic models (PBPK). This is known as in vitro to in vivo extrapolation (IVIVE).

4. Modeling softwares that can perform sophisticated species scaling (prediction of human parameters from nonclinical species such as rat, dog, monkey, etc.)

5. Risk assessment and exposure modeling tools that enable calculation of hazard and risk in populations on a whole and specific subpopulations.

Toxicity Testing in Chemicals vs. Pharmaceuticals

Regulatory based toxicity testing is different for pharmaceutical based compounds versus chemical and agrochemical compounds. Non clinical safety assessment for phamaceutical products is governed by the different stages of drug development while the toxicological data requirement for chemicals/agrochemicals is based on the amount (tonnage) produced. While most pharmaceutical compounds undergo extensive animal toxicity testing, there are thousands of chemicals in commerce today that have undergone very limited or non toxicity testing. In order to address this several governmental mandates are being put into action.

Government Initiatives & Mandates for 21st Century Toxicity Testing

- In Europe, the Registration Evaluation Authorization and Restriction of Chemicals (REACH) was initially implemented in 2007.
- This substantially altered the safety testing performed on new as well as existing chemicals.
- In The United States too several initiatives to increase safety testing on more and more chemicals are underway which would increase the cost of safety assessments astronomically.
- The regulations under REACH have been directly estimated to cost the industry more than 4.2 billion dollars (Brown, 2003).

Figure 2.6.4.3.1: Brown, V. J. (2003). REACHing for chemical safety. Environ. Health Perspect. 111, A766–A769.

21st Century Toxicity Testing

Traditional toxicity testing involves the use of a lot of animals and is an extremely expensive and time consuming process. In order to address the large number of untested chemicals the US Environmental Protection Agency (EPA) initiated the **ToxCast program.** The ToxCast Program is a high-throughput screening program that would enable the prioritization of chemicals so that resources can be channelized towards those chemicals that possess the greatest risk to human safety.

THE TOXCAST PROGRAM



The ToxCast program developed and utilized automated in vitro assays (to test effects of chemicals on various biological processes using living cells, isolated proteins etc.). The assay designs included endpoints such as cytotoxicity, enzyme activity, endocrine endpoints, gene expression etc. A total of approximately 600 endpoints were evaluated.

The biggest question to answer was how relevant was this data to human safety and how to utilize this *in vitro* information to predict human safety.

Figure 2.6.4.3.2: National Institute of Environmental Health Sciences – NIH, 50th Anniversary

The Overall Perspective

In vitro assays may be relevant in the light of the fact that overall disesase/toxicology processes are actually mediated by molecular and cellular perturbations. However, the overall picture at the whole organism level includes several other complex factors such as pharmacokinectics of the compound, metabolism, clearance etc. This gap could be filled by utilizing computational modeling tools that would utilize the in vitro data and integrate them into human physiology with the help of pharmacokinetic (PK) or physiologically based pharmacokinetic models (PBPK).

Understanding the concept of effects Vs. adverse Effects

In order to conduct safety assessments it is important to understand the concept of “adverse effects” versus simply “effects”. It is also important to understand the significance of biological relevance of isolated *in vitro*, molecular assays as it pertains to the whole organism.

Adverse Outcome Pathways (AOPs) have been developed to try and link the causal molecular initiating event to a host of intermediate processes at the cellular level that finally leads to adverse outcome (AO) in the whole organism that can be used for safety assessment purposes.

Adverse Outcome Pathways

The AOP programme was launched by the Organization for Economic Co-operation and Development (OECD) in 2012. The objective was to link the main molecular initiating event with the phenotypic/functional toxicity/adverse effect at the organism level.



Figure 2.6.4.3.3: Adverse Outcome Pathways

Using Systems Toxicology Approach For Toxicity Predictions

The molecular initiating events could be used in selecting in vitro assays that could have possible potential for predicting toxicity at the whole organism level.

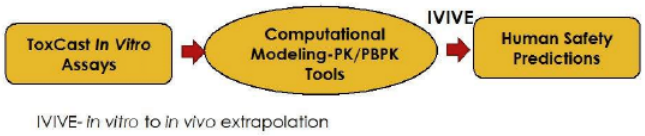


Figure 2.6.4.3.4: IVIVE- in vitro to in vivo extrapolation

Topic 3: Key Points

In this section, we explored the following main points:

- 1: Tools and Technologies, used in the field of systems toxicology.

- 2: Concept of Adverse effects and AOPs .

Knowledge Check

1. The ToxCast program utilizes which kind of assays?

In vivo

In silico

In vitro

None of the above

Answer

in vitro

2. Adverse Outcome Pathway (AOP) Programme was launched by...

National Toxicology Program (NTP)

National Institute of Health (NIH)

Environmental protection Agency (EPA)

Organization for Economic Co-Operation and Development

Answer

Organization for Economic Co-Operation and Development

This page titled 2.6.4.3: Tools and Technologies in Systems Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **4.3: Tools and Technologies in Systems Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.4.4: Other Approaches for Predictive Toxicity Modeling

Learning Objectives

- 1: Understand what Quantitative Structure Activity Relationships are and how they are used in the field of systems toxicology.

Quantitative Structure Activity Relationships (QSAR)

QSAR models are classification models that are used to link particular structures in some molecules to causative/adverse effects at cellular/organism level. This is based on the fact that compounds that are structurally similar would have similar mechanisms of mediating toxicity. For example compounds belonging to the structural class of triazoles have been reported to cause developmental malformations such as cleft palate in rats.

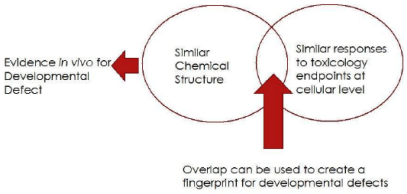
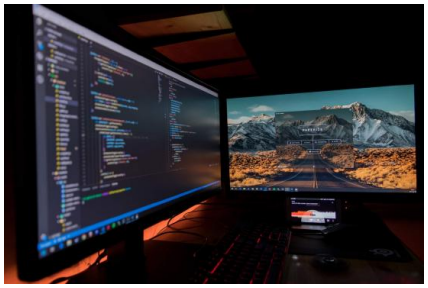


Figure 2.6.4.4.1: QSAR Approaches for Predictive Toxicity Testing

Mechanisms at the cellular level include events such as positive response for the TGFb1 signaling pathway. This information can be used to build a molecular fingerprint where compound having similar chemical structures as triazoles and having a positive response to TGFb1 signaling pathway could be flagged for potential developmental toxicity (teratogens).

QSAR Approaches for Predictive Toxicity Testing



This method however requires a lot of data in order to build robust databases. This also requires validation of QSAR relationships with large sets of *in vitro* and *in vivo* datasets. This requires extensive collaboration between scientists across different disciplines of toxicology.

Currently, several softwares are available commercially that are used for predictive toxicology assessments. These softwares have been validated against compounds across different classes and various toxicological end points. Utilizing such tools requires detailed research and information on the extent of evaluation and validation conducted in order to develop these tools. For example softwares that have been validated with large data sets for pharmaceutical compounds may not be appropriate for using in a chemical space and vice versa.

Names of certain commercially available Computational Tools Used In Industry

- QSAR tools: Derek, GastroPlus, ADMET Predictor, OECD QSAR Toolbox, ACD ToxSuite
- PK/PBPK Modeling Tools: GastroPlus, Simcyp, WinNonlin, Berkeley Madonna

Topic 4: Key Points

In this section, we explored the following main points:

- 1: The application of QSARS in systems toxicology.

Knowledge Check

1. QSAR databases link:

Relationship between chemical structures of compounds and their activity/toxicity

Relationship between different responses at different organizational levels

Physiologically relevant modeling tools

In vitro to in vivo extrapolation

Answer

Relationship between chemical structures of compounds and their activity/toxicity

2. Development of efficient QSAR tools depend on...

The user’s familiarity with the software/tool

Knowledge of multiple QSAR models

A large database of validation training sets

All of the above

Answer

A large database of validation training sets

This page titled 2.6.4.4: Other Approaches for Predictive Toxicity Modeling is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **4.4: Other Approaches for Predictive Toxicity Modeling** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.4.5: Technologies Used In Systems Biology/Toxicology

Learning Objectives

- 1: Understand the different omics technologies and how they are used.

"Systems biology/toxicology uses very powerful high-throughput platforms/tools such as the “omics” technologies. The human genome was first sequenced in 2003. It took 13 years and was extremely expensive. Advancement in technology has now made it possible to have genome sequencing completed in less than a day. It has also made handling and interpretation of large volumes of data at different levels (gene, transcriptome, protein, metabolic) possible. Scientists across different disciplines are now using these technologies to provide a more holistic approach to disease and toxicity."

Systems Biology in Toxicology and Environmental Health; First Edition, Chapter 1

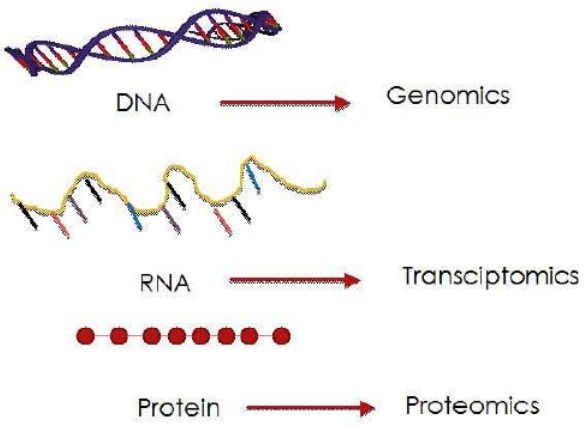


Figure 2.6.4.5.1: A schematic approach for “omics” technologies.

Parts of illustration from www.pixabay.com

Technologies Used In Systems Biology/Toxicology

Genomics: This refers to the technology which allows us to study the complete genetic material of an organism. This involves DNA sequencing and analysis. Some of the most common methods used for high-throughput genotyping for genome wide association studies (GWAS) include Illumina Omni Arrays which can simultaneously analyze upto 5 million markers per sample. Other examples are the HumanOmni 5 Quad (Omni 5) and the Affymetrix platforms.

GWAS: A GWAS involves analysis of genetic variants in different individuals to study particular traits associated with genetic variations. This enables scientists to study the underlying mechanisms of different diseases at the genomic level. Single base pair changes are the most common form of variants of the genome and are known as single nucleotide polymorphisms (SNP). While most of them are functionally harmless, rare ones can lead to changes at the protein level leading to functional impairment and diseases.

Genomics: Example Application

SNPs that results in functional changes at the protein level thereby causing a difference in phenotype such as a disease state is known as a mutation. These are considered rare genetic variants. An example of a disease state is cystic fibrosis which is cause by multiple mutation in the CFTR gene. This was made possible by genotyping families that were affected by this disease and identifying markers (genetic variants) that could be linked to this disease. This is known as linkage analysis. This type of analysis was also successfully used to identify unique mutations that lead to rare diseases like Huntington’s disease. While this approach has been successful for rare diseases it has been a challenge to utilize this for more common disease states such as cancer, heart and liver disease etc.

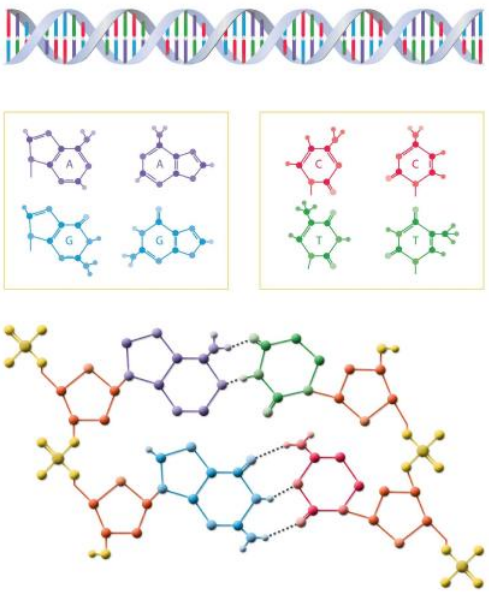


Figure \PageInde2: PLoS Comput Biol. 2012 Dec; 8(12)

Transcriptomics

"Transcriptomics is the study of all transcriptomes (RNA) in a genome. In other words, it is the study that enables us to study large volumes (> 200,000) of data related to gene expression at the RNA level.

Traditionally RNA expression was studied through a technique known as Northern blots. But this technique could not handle large volumes of data. Current techniques used for high-throughput data analysis are different kinds of microarrays and biochips. The most commonly used are the DNA based microarrays.

Quantitative gene expression analysis helps us understand the difference in expression of genetic product between different cells,tissues, species etc. Such microarrays are commercially made available from companies like Affymetrix, Agilent, Applied Microarrays, Illumina etc."

Systems Biology in Toxicology and Environmental Health; First Edition, Chapter 4

Different kinds of microarrays

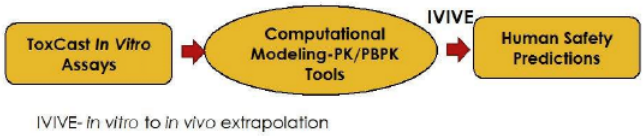
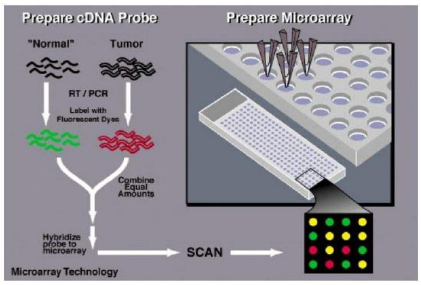
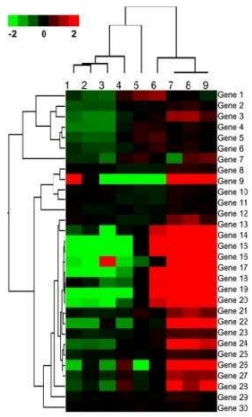


Figure 2.6.4.5.3: Different kinds of microarrays. Rebecca C. Fry. Systems Biology in Toxicology and Environmental Health, Chapter 4 (Kindle Location 1954). Elsevier Inc.

Analysis of microarray data



Microarray data provides relative expression levels of various genes. Quantitaive analysis of microarray data require sophisticated statistical tools and is expressed in the form of a “heat map” which provides expression patterns (upregulation/downregulation of genes) in the form of various colors. Significance analysis of microarrays (SAM is common technique) is a statistical tools that allows for quantitative determination of expression patterns.Figure 2.6.4.5.4: Sample processing for microarrays. https://www.genome.gov/10000533/dna-...ay-technology/ The output which is the heatmap has different colors signifying differential expression patterns. In the example given below red color signifies upregulated gene expression while green signifies downregulation.Figure 2.6.4.5.5: Heatmap generated from microarray. Rebecca C. Fry. Systems Biology in Toxicology and Environmental Health, Chapter 4 (Kindle Location 1954). Elsevier Inc.



Proteomics

"Proteomics is the field that studies total proteins in a system. This involves high-throughput profiling of proteins. This kind of expression pattern is especially useful since it allows us to study post translational effects, since all transcription products are not always converted to proteins; however, all functional aspects at the phenotypic level are almost always driven by proteins.

Proteomics is a powerful tool that can help us in identification of protein biomarkers specific to toxicity due to particular exposures or specific diseases states. Mass spectrometers are commercially sold by several companies such as Waters, Thermofisher, Agilent, Shimadzu, Perkin Elmer etc.

A mass spectrometer is used for analysis of proteomics data. Interpretation requires sophisticated bioinformatics tools."

Rebecca C. Fry. Systems Biology in Toxicology and Environmental Health, Chapter 4 (Kindle Location 1954). Elsevier Inc.

Sample processing and workflow for proteomics studies:

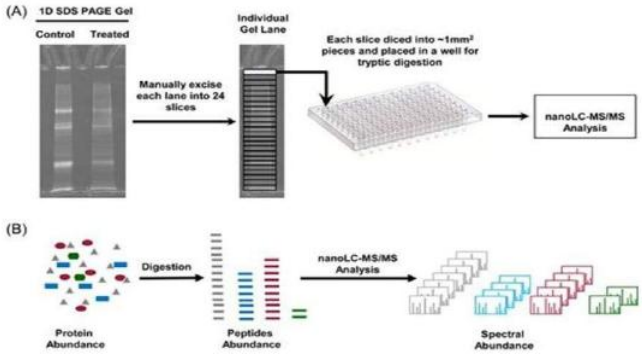


Figure 2.6.4.5.6: Sample processing and workflow for proteomics studies. Rebecca C. Fry. Systems Biology in Toxicology and Environmental Health, Chapter 4 (Kindle Location 1954). Elsevier Inc.

Metabolomics

"Metabolomics refers to the study of metabolites (low molecular weight products of cellular /biological processes that are found in cells, tissues, biological fluids etc.). Metabolomics provides an understanding of the differences in the biochemistry between different variants (such as diseased and healthy patient populations, control group versus treatment group, high dose group versus low dose group etc.)

Metabolic profiling can be easily performed in biological matrices such as urine, blood, plasma, serum and also a wide variety of tissues. Hence this can be used very efficiently in human health and safety assessment for evaluating biomarkers for exposure to toxicity to various agents. It can be used in drug discovery in the pharmaceutical industry for comparing metabolic profile between different dose level and control groups. In the clinical setting metabolomics can provide an understanding of differences in metabolite products in diseased versus healthy patient populations.

Platforms/tools used in metabolomics is similar to proteomics. Liquid chromatography mass spectrometers (MS) are broadly used for metabolite profiling of different biological matrices. Additionally, nuclear magnetic resonance (NMS) imaging is also widely used for metabolomics analysis. A major difference between the two tools is that NMR is a non destructive process and samples used for NMR can be reused for other purposes or returned to the biorepository whereas this is not possible with the LCMS method.

Metabolomic studies can be conducted using a “targeted approach” where a few selected analytes (metabolites) are analyzed based on hypothetical research. Such studies are mainly performed in early discovery phases. Alternatively, a more “broad spectrum” approach may be used in order to develop biomarkers for certain treatments or specific disease states.

As with other “omics” approaches, state-of-the art bioinformatics and statistical tools are used for quantitative interpretation of the data generated from metabolomics platforms (LCMS, NMR)."

Rebecca C. Fry. Systems Biology in Toxicology and Environmental Health, Chapter 4 (Kindle Location 1954). Elsevier Inc.

Metabolomics Workflow

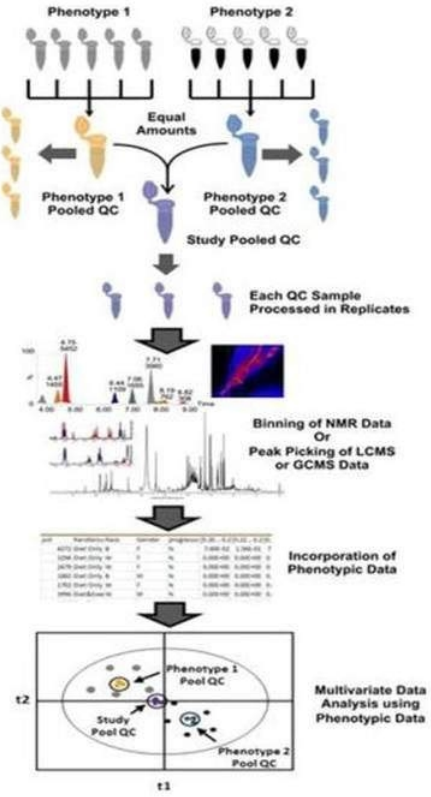


Figure 2.6.4.5.7: Metabolomics Workflow. Rebecca C. Fry. Systems Biology in Toxicology and Environmental Health, Chapter 4 (Kindle Location 1954). Elsevier Inc.

Topic 5: Key Points

In this section, we explored the following main points:

- 1: The application of the different “omics” technologies in systems toxicology.

Knowledge Check

1. Which of the following processes is a non destructive process in which samples can reused or returned back to the biorepository?

Microarray

DNA sequencing

NMR

Protein Isolation

Answer

NMR

2. Which of the following technologies is used to study complete RNA profiles

Nuclear Magnetic Resonance

Microarrays

DNA sequencing

Physiologically based pharmacokinetic models

Answer

Microarrays

3. The central dogma of molecular biology states is...

DNA to RNA and RNA to protein

RNA to protein to DNA

Protein to DNA to RNA

None of the above

Answer

DNA to RNA and RNA to protein

4. Liquid Chromatography mass spectroscopy is used for...

Genomics

Proteomics

Metabolomics

Proteomics and metabolomics

Answer

Proteomics and metabolomics

This page titled 2.6.4.5: Technologies Used In Systems Biology/Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **4.5: Technologies Used In Systems Biology/Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.4.6: Takeaways Summary

The tremendous advancement at the technological level has made it possible to generate data at “high-throughput” levels. It has also enabled scientists to study toxicological processes from a more holistic approach. Instead of answering single questions at a time, a more comprehensive approach is now being applied to understand toxicological responses. “Systems” approach is being effectively used in industry, government and clinical settings.

In biopharmaceutical/chemical industries, thousands of molecules are screened in the early discovery phase to select target compounds with efficacy. Similarly, compounds can also be screened based on their toxicity fingerprints. Molecular fingerprints generated for each class of compounds/ specific chemistries can be used to screen compounds for different indications for future uses. This has also significantly reduced the time and increased efficiency in discovery programs in industry settings.

The government has also been using “the systems” approach successfully for safety assessment programs. Efforts such as the ToxCast and Tox 21 are examples where “systems” toxicology is being efficiently used to prioritize animal testing towards chemicals that pose the greatest risk to human health and safety.

In the clinical setting more and more information is becoming available via the “systems” approach for specific disease states that enable physicians and researchers to develop personalized medicines based on specific needs of patients.

While, all these efforts mark the beginning of a very promising future, there is still a lot of research necessary to utilize these tools and technologies to their full potential.

REFERENCES

- Systems Biology in Toxicology and Environmantal Health. First Edition. Editor Rebecca Fry. Chapters 1 and 4

- Incorporating Human Dosimetry and Exposure into High-Throughput In Vitro Toxicity Screening. *Rotroff et al., Toxicological Sciences;* 117(2), 348–358 (2010)

This page titled 2.6.4.6: Takeaways Summary is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **4.6: Takeaways Summary** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.4.7: Section 4 Final Evaluation

1. *In vitro* to *in vivo* extrapolation (IVIVE) is modeling software that can utilize *in vitro* data and mathematically translate that into relevant *in vivo* information utilizing pharmacokinetic (PK).

True

False

Answer

True

2. Which of the following is a common tool used for both Quantitative Structure Activity Relationships (QSAR) and Physiologically based pharmacokinetic (PBPK)?

GastroPlus

ADMET Predictor

ACD Toxsuite

Derek

Answer

GastroPlus

3. HumanOmni 5 Quad (Omni 5), illumina Omni Arrays and the Affymetrix platforms are common methods used in:

Transcriptomics

Genomics

Proteomics

Metabolomics

Answer

Genomics

4. Quantitative Structure Activity Relationships (QSAR) approaches for predictive toxicity testing requires all the following EXCEPT:

Lots of data in order to build robust databases.

Validation of QSAR relationships with large sets of in vitro and in vivo datasets.

Extensive collaboration between scientists across different disciplines of toxicology.

A mass spectrophotometer.

Answer

A mass spectrophotometer.

5. Computational modeling tools utilizes the in vitro data and integrate them into human physiology with the help of...

Physiologically based pharmacokinetic model.

High-throughput in vitro model.

Hypothesis and diagnostic model.

Mathematical and computational model.

A

Answer

Physiologically based pharmacokinetic model.

6. Systems toxicology aims to fill this gap and utilize these data from different systems and integrate them into meaningful assessment for safety. This gap is:

Lack of modern equipment and technology for data collection.

Lack of interpretation and utilization of collected data.

Lack of human resources and management.

Lack of organ and system models.

Answer

Lack of interpretation and utilization of collected data.

7. Adverse Output Pathways (AOP) can be assessed in the following stages:

Molecular initiating event

Intermediate response at cellular level

Toxic response at organic level

All of the above

Answer

All of the above

8. In Metabolomics, the major difference between liquid chromatography mass spectrometers (LCMS) and nuclear magnetic resonance (NMS) is...

NMR samples cannot be reused for other purposes while LCMS samples can be reused for other purposes.

NMR is a destructive process whereas LCMS is a non-destructive process.

NMR is a non-destructive process whereas LCMS is a destructive process.

NMR samples cannot be returned to the biorepository whereas LCMS samples can be returned to the biorepository.

Answer

NMR is a non-destructive process whereas LCMS is a destructive process.

9. In the study of the underlying mechanisms of different diseases at the genomic level, the most common form of variants at this level is:

Polyploidy

Aneuploidy

Single Nucleotide Polymorphism (SNP)

Micronuclei variants

Answer

Single Nucleotide Polymorphism (SNP)

10. ____ is used for analysis of proteomics data:

A mass spectrophotometer

DNA micro array

Illumni Omni Array

DNA sequencing

Answer

A mass spectrophotometer

11. The ToxCast program utilizes which kind of assays?

In vivo

In vitro

In silico

All of the above

Answer

In vitro

12. The following agencies were established to address the challenge of large number of chemicals that go to market with very limited or no toxicity testing:

Environmental Protection Agency (EPA) in the US.

Registration Evaluation Authorization and Restriction of Chemicals (REACH) in Europe.

Both A & B are correct.

Occupational Safety and Health Administration (OSHA).

Answer

Both A & B are correct.

This page titled 2.6.4.7: Section 4 Final Evaluation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **Section 4 Final Evaluation** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.5: Regulatory Toxicology

Regulatory toxicology is where the science of toxicology meets the regulations, policies and guidelines that protect human health and the environment from chemicals. In this e-module you will learn about global, regional, national, state, and non-governmental regulatory toxicology.

2.6.5.1: Introduction to Regulatory Toxicology

2.6.5.2: Global Regulatory Toxicology

2.6.5.3: Topic 3- Regional Regulatory Toxicology

2.6.5.4: National Regulatory Toxicology

2.6.5.5: State Regulatory Toxicology

2.6.5.6: Non-Governmental Regulatory Toxicology

2.6.5.7: Section 5 Final Evaluation

This page titled 2.6.5: Regulatory Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.5.1: Introduction to Regulatory Toxicology

Learning Objectives

- 1: Identify regulatory toxicology inside and outside of government.
- 2: Give examples of regulatory toxicology at various scales and locales.
- 3: Explain the difference between regulation and guidance.

What is Regulatory Toxicology?

Regulatory Toxicology is where the science of toxicology meets the regulations, policies and guidelines that protect human health and the environment from chemicals. Regulatory toxicology commonly is associated with government agencies. These agencies may vary dramatically in their size and scope. For example, the United Nations covers the entire globe, while agencies within a city are limited to the area covered by the municipality.

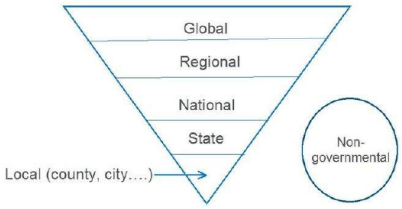


Figure 2.6.5.1.1: Introduction to regulatory toxicology

Regulatory agencies generally have specific focus areas that they address. For example, the U.S. Occupational Safety and Health Administration (OSHA) covers hazardous chemicals in the workplace, while the U.S. Consumer Product Safety Commission (CPSC) addresses chemical hazards in consumer products. Lastly, regulatory toxicology also occurs in non-governmental agencies such as professional societies, private industry and various advocacy groups. Further discussion of these is provided later in the module. More information on OSHA and CPS.

What is the difference between a regulation, policy, and guideline?

A regulation is a rule or order issued by a governmental authority that has the force of law. Often regulations are developed by experts in a governmental authority to enforce legislation. An example of a regulation is the Food Quality Protection Act (FQPA) passed by the U.S. Congress and signed into law by the President in 1996.

Policies and guidelines are principles and approaches that clarify and interpret regulations. As such, policies and guidelines do not carry the force of law but provide important direction.

The Federal Insecticide, Fungicide and Rodenticide Act (FIFRA) is a U.S. regulation that regulates the broad class of chemicals used as pesticides (i.e., substances used to combat “pests”). The U.S. Environmental Protection Agency (EPA) has authority over FIFRA, and has in turn established many policies and guidelines concerning pesticides.

One area that EPA has established multiple policies and guidelines is in the registration of pesticides (i.e., EPA review and approval).

Topic 1: Key Points

In this section, we explored the following main points:

- 1: What is Regulatory Toxicology?
- 2: What is the difference between a regulation, policy and guidance

2. What is the difference between a regulation, policy and guidance

Knowledge Check

1. What is regulatory toxicology?

The forces that govern the regulation of homeostasis in the body

The branch of toxicology that studies regulations

The intersection of the science of toxicology with the regulatory world protecting human health and the environment

Regulations that dictate the science underlying toxicology

Answer

The intersection of the science of toxicology with the regulatory world protecting human health and the environment

2. Which of the following have the force of law?

A regulation

A policy

A guideline

All of the above

Answer

A regulation

This page titled 2.6.5.1: Introduction to Regulatory Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **5.1: Introduction to Regulatory Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.5.2: Global Regulatory Toxicology

Learning Objectives

- 1: Define what is meant by “global regulatory toxicology”
- 2: Give an example of a global regulatory guideline

What is Global Regulatory Toxicology?

Global regulatory toxicology is exactly as it sounds, it deals with regulatory toxicology on a global scale (i.e., the entire planet). Relative to other jurisdictions (e.g., nations, states, cities), there are few regulatory toxicology initiatives that are global in nature. Some initiatives (e.g., clean drinking water, clean air) with similar intent may span many parts of the globe and appear global, but they lack a global consensus.

Global regulatory toxicology initiatives often originate from activities by the United Nations (UN), a global organization bringing together member countries to confront common challenges.

Example of Global Regulatory Toxicology: Globally Harmonized System (GHS)

Full name is the Globally Harmonized System of Classification and Labelling of Chemicals (GHS). It was created by the UN. Work on GHS began in 1992 and the first edition was released in 2003. GHS is updated every two years. The goal is to harmonize the criteria by which chemicals are classified in terms of their hazards. Hazards include physical (e.g., flammability), environmental (e.g., toxicity to fish), and human heath (e.g., acute toxicity people).

Example of GHS classification – Hazards to the Eye

Category	Hazard	Hazard statement	Requirement for eye protection?	Signal Word	Symbol
1	Irreversible effects	Causes serious eye damage.	Yes	DANGER!	
2A	Reversible effects	Causes serious eye irritation.	Yes	WARNING!	
2B	Reversible effects	Causes eye irritation.	No	WARNING!	None
Not classified	None	None	No	None	None

Figure 2.6.5.2.1: The Globally Harmonized System (GHS)

The Globally Harmonized System (GHS)

Prior to Globally Harmonized System (GHS), each country had its own criteria for hazard classification, and some countries had multiple criteria. This provided challenges and confusion to the general public and other stakeholders. GHS is a guideline, not a regulation. However, once adopted by a country, GHS generally becomes a regulation.

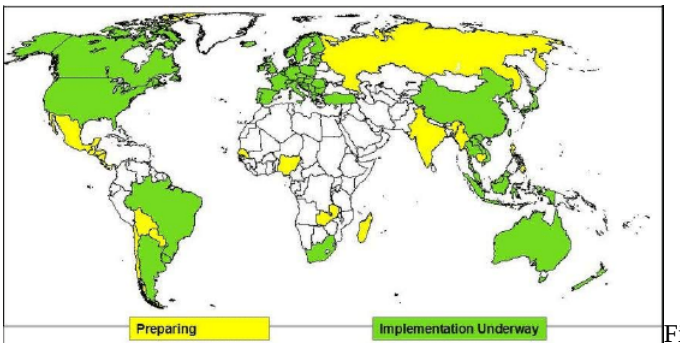


Figure 2.6.5.2.2: GHS Implementation Status – May, 2017

Topic 2: Key Points

In this section, we explored the following main points:

- 1: What is Global Regulatory Toxicology?

- 2: The Globally Harmonized System (GHS)

Knowledge Check

1. Which of the following generally is involved with establishing a global regulatory toxicology initiative?

Environmental Protection Agency (EPA)

United Nations (UN)

Food and Drug Administration (FDA)

All of the above

Answer

United Nations (UN)

2. Which of the following is an example of a global regulatory toxicology initiative?

Clean Air Act

Clean Water Act

Globally Harmonized System

Safe Drinking Water Act

Answer

Globally Harmonized System

This page titled 2.6.5.2: Global Regulatory Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **5.2: Global Regulatory Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.5.3: Topic 3- Regional Regulatory Toxicology

Learning Objectives

- 1: Define what is meant by a “regional regulatory toxicology”
- 2: Give an example of a regional regulation

What is Regional Regulatory Toxicology?

Regional regulatory toxicology exists between global and national regulatory toxicology. It deals with regulatory toxicology that includes multiple countries. Often these countries are adjacent or in close proximity (e.g., United States and Canada) but that is not a requirement.

A group of nations that frequently exert regulations on a regional scale is the European Union (EU). The EU started in 1951 with six European countries agreeing to cooperate and has expanded over the ensuing decades to the current list of 28 member countries. -One current member of the EU, the United Kingdom has decided through a popular vote to eventually exit the EU (i.e., Brexit).

It started as a paper in 2001 and entered into force on June 1, 2007. It's a new approach for the regulation of chemicals in the EU. Required key hazard information to be obtained for existing chemicals to remain in commerce, or for a new chemicals to enter into commerce.

Information falls into three broad hazard categories: physical-chemical, human health and environmental health. The types and amounts of information varies depending on the volume of chemical used in commerce. The more the chemical is used, the more the data.



In addition to hazard information, REACH requires assessments to determine if chemicals in commerce pose unreasonable risk(s). REACH has resulted in the assessment and/or re-assessment of 1000s of chemicals. A large database of toxicity information was created. Some countries are using REACH as a model to develop new, or update existing, chemical regulations, for example, South Korea with K-REACH.

Topic 3: Key Points

In this section, we explored the following main points:

- 1: What is Regional Regulatory Toxicology?
- 2: The Registration, Evaluation, Authorization and restriction of CHemicals (REACH) regulation

Knowledge Check

1. Which of the following generally is a trait commonly found in a regional regulatory toxicology initiative?

Involvement of multiple regulatory agencies within a country

Involvement of all/most countries around the globe

Involvement of multiple states within the United States

Involvement of two or more countries in a regulatory initiative

Answer

Involvement of two or more countries in a regulatory initiative

2. Which of the following is an example of a regional regulatory toxicology initiative?

GHS

REACH

K-REACH

FQPA

Answer

REACH

This page titled 2.6.5.3: Topic 3- Regional Regulatory Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **5.3: Topic 3: Regional Regulatory Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.5.4: National Regulatory Toxicology

Learning Objectives

After completing this lesson, you will be able to:

- Define what is meant by “National Regulatory Toxicology”.
- Give an example of a national regulatory regulation.

What is National Regulatory Toxicology?

As the name suggests, it focuses on regulatory toxicology at the national level (e.g., United States, Canada, China). Regulations are set by authorities that apply for the entire country. The focus on the United States (US) deals with regulations established at the federal level, and all states within the US must comply with federal regulations.

Agency	Population Served	Mainly Regulates Substances in:
EPA	General population (humans); Environment (e.g., aquatic species)	• Environment (e.g., air, soil, water)
FDA	General population; Medical Patients	• Food • Drugs and Medical Devices
OSHA	Workers	• Occupational environment (e.g., air in a factory)
CPSC	Consumers	• Consumer products (e.g., shampoos, toys, clothing)

Figure 2.6.5.4.1: Example agencies that set national toxicology regulations in the U.S.

Example agencies that set national toxicology regulations in the U.S.:

- Environmental Protection Agency (EPA)
- Food and Drug Administration (FDA)
- Occupational Safety and Health Administration (OSHA)
- Consumer Product Safety Commission (CPSC)

All agencies aim to set regulations to protect against adverse health effects from exposure to substances. Different agencies regulate on different types of substances, that are used for different purposes and affect different populations.

Goal of **EPA** is to protect human and environmental health; deals with exposures to substances in the environment (e.g., air, water, soil). It typically pertains to exposures that the general population generally does not have control over (e.g., air pollution) as well as modernizing chemical regulation in the US, e.g. Toxic Substances Control Act (TSCA) was recently revised.

TSCA was first passed in 1976. TSCA was revised with passage of the **Frank R. Lautenberg Chemical Safety for the 21st Century Act** in 2016. It regulates new or already existing chemicals manufactured or in use in the US.

Topic 4: Key Points

In this section, we explored the following main points:

- 1: What is National Regulatory Toxicology?
- 2: Examples of National Agencies in the U.S.
- 3: What do these Agencies regulate?
- 4: Highlight: EPA and TSCA Reform.

Knowledge Check

1. EPA regulates chemicals that are in:

Food

Environmental media (e.g., water, air)

Drugs and Medical Devices

Shampoo

Answer

Environmental media (e.g., water, air)

2.US States can choose not to comply with regulations set by the FDA

True

False

Answer

False

This page titled 2.6.5.4: National Regulatory Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **5.4: National Regulatory Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.5.5: State Regulatory Toxicology

Learning Objectives

- 1: Define what is meant by “State Regulatory Toxicology”.
- 2: Give an example of a state regulatory regulation.

What is State Regulatory Toxicology?

As the name suggests, it deals with regulatory toxicology for the US on a state level (e.g. Texas, Delaware, Arizona). It was established by authorities applicable to a given state. State regulatory toxicology only applies with the state’s boundaries; however, they may influence adjacent states or even national regulatory toxicology.

State regulatory toxicology is under the layer of the US national government. It sits under a complex web of state and local laws and policies, in addition to regulatory authorities. The make-up of state and local governments varies widely across the US; while they have mutual specific features, their organizations differ. Whatever their design, state and local governments can sometimes have a much greater impact on people's lives than the federal government.

The Federal-State Toxicology and Risk Analysis Committee (FSTRAC) is made up of representatives from U.S. state health and environmental agencies and U.S. EPA personnel.

FSTRAC is an integral part of EPA’s communication strategy with states and tribes for human health risks associated with water contamination. It fosters cooperation, consistency, and an understanding of EPA’s and different states’ goals and problems in human health risk assessment. Additionally, it allows states and the federal government to work together on issues related to the development and implementation of regulations and criteria under the Safe Drinking Water Act and Clean Water Act.

FSTRAC members have supported development of Human Health Benchmarks for Pesticides (HHBP).

- 1: Represent levels of pesticides in drinking water that are not anticipated to cause health effects.
- 2: Used to help assess drinking water quality for pesticides that do not have other regulatory toxicology standards.

Examples of agencies that set national toxicology regulations in the U.S.:

The goal of state agencies is the same as federal agencies, protect people and the environment from health effects associated with chemical exposures.

States have differing regulatory toxicology requirements and focuses. Reasons for the differences could be due to: state history, geography, culture, population size and diversity, major industries, etc…

Examples of state regulatory toxicology as it concerns chemicals in air:

TEXAS

Effects Screening Levels (ESLs) – Air concentrations generally applicable to a specific chemical set by the Texas Commission on Environmental Quality (TCEQ). Play an important role in the regulation of air emissions from companies located in the state. List of ESLs here.

CALIFORNIA

Reference Exposure Levels (RELs) – Air concentrations generally applicable to a specific chemical set by the California Environmental Protection Agency (CalEPA). Represent an air concentration that does not pose a health risk to people. List of RELs here.

Agency	State Specific Rule or Regulation	Requirement
CalEPA	Proposition 65	• List of chemicals known to the state to cause cancer or reproductive toxicity
TCEQ	Texas Risk Reduction Program (TRRP)	• Established cleanup standards for contaminated chemical sites

Figure 2.6.5.5.1: Example of state toxicology regulations and the requirement that must be followed.

Example agencies that set state toxicology regulations

Agency	Regulation
Office of Environmental Health Hazard Assessment	Proposition 65 (Prop 65)
State of Washington Department of Ecology	Children's Safe Products Act (CSPA)
Minnesota Department of Health – Environmental Health	Minnesota Wellhead Protection Program

Figure 2.6.5.5.2: Example agencies that set state toxicology regulations.

An example of a State Regulatory Agency is the **CalEPA Office of Environmental Health Hazard Assessment (OEHHA)**: Provides requirements and guidance for the California Proposition 65 Safe Drinking Water and Toxic Enforcement Act of 1986 (Proposition 65).

Proposition 65

Purpose

Enable consumers to make informed decisions regarding chemical exposures.

Reason

Established to protect California citizens from chemicals known to the state to cause cancer, birth defects, or other reproductive harms.

Scope

Addresses chemical exposures to the citizens of California that may occur through consumer products, workplace exposures, and exposures occurring via the environment.

Basics

OEHHA publishes a list of chemicals known to cause cancer, birth defects or other reproductive harm. The list is updated regularly and currently contains approximately 900 chemicals. Once a chemical is listed, companies have 12 months to comply with warning requirements under the regulation.

Proposition 65 is referred to as a "risk-based" regulation in that the warning requirements only apply if the risk from chemical exposure are too high as defined by the regulation.

Exposure examples:

- Oral
- Inhalation
- Skin contact

Proposition 65 upcoming changes…

The regulation has undergone revisions: New Proposition 65 Warnings, that will now require companies to add a symbol and change the phrasing of the warning. For example: "**WARNING:** This product can expose you to chemicals including arsenic, which is known to the State of California to cause cancer. For more information, visit here.

Topic 5: Key Points

In this section, we explored the following main points:

- 1: What is State Regulatory Toxicology?
- 2: An example of how federal and state agencies work together
  - What is the Federal-State Toxicology and Risk Analysis Committee (FSTRAC)
  - Two example outcomes of the FSTRAC's workgroup
- 3: Example of State Agencies
- 4: Highlight two state specific guidance and rules as it concerns chemicals in air
  - Guidance
    - Effects Screening Levels (ESLs)
    - Reference Exposure Levels (RELs)
- 5: Rules
  - California Proposition 65
  - Texas Risk Reduction Program rule

Knowledge Check

1. Proposition 65 is a regulation established in:

Florida

Texas

California

New York

Answer

California

2. Which of the following could explain some of the variability observed between states in terms of regulatory toxicology?

Degree of urbanization

Type of natural resources

Degree of industrialization

All of the above

Answer

All of the above

This page titled 2.6.5.5: State Regulatory Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **5.5: State Regulatory Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.5.6: Non-Governmental Regulatory Toxicology

Learning Objectives

- 1: Define what is meant by a “Non-Governmental Regulatory Toxicology”.
- 2: Give several examples of non-governmental entities that conduct and influence regulatory toxicology.

What is Non-Governmental Regulatory Toxicology?

They are groups that are not officially part of a government agency that conduct and/or influence regulatory toxicology. Examples of such groups may include not-for-profit organizations, advocacy groups, professional societies, industry trade associations and individual companies. The term “NGO” (short for non-governmental organization(s)) is sometimes used; however, the term “NGO” often is limited to advocacy groups.

These groups often can conduct or influence regulatory toxicology on a much faster timeframe than government agencies. Government agencies often take many months or years to enact regulations, policies, and guidance; while non-governmental entities often proceed at a fraction of this time. Lastly, they often lack the transparency found in government agencies in terms of decision making, as well as the ability to provide input.

Examples of a non-governmental entities that conduct and influence regulatory toxicology

- Advocacy groups – Often have a particular focus such as the proper treatment of animals, environmental sustainability, women’s health…
- Individual companies – For-profit businesses that decide to address a particular topic related to regulatory toxicology. Examples include eliminating use of certain chemicals, reduce use of water, increasing recycling…
- Trade Associations – represent business that have a common or similar commercial interest (e.g., chemical manufacture, consumer products). Often create industry best practices or guidance for issues related to regulatory toxicology.

Topic 6: Key Points

In this section, we explored the following main points:

- 1: What is Non-Governmental Regulatory Toxicology?
- 2: Examples of different types of non-governmental entities that conduct and influence regulatory toxicology.

Knowledge Check

1. Which of the following is a trait found in a non-governmental entities?

Must engage with the general public prior to taking actions.

Have lengthy processes that often take decades to complete.

Is not a government agency or formally associated with a government agency.

A narrowly defined entity that is not common in society.

Answer

Is not a government agency or formally associated with a government agency.

2. Which of the following is an example of non-governmental regulatory toxicology?

A not-for profit company developing standards and conducting certifications regarding the sustainability of products.

An advocacy group that provides input to government and industry related to reducing lead in consumer products.

An organization representing chemists develops guidance on workplace safety concerning corrosive chemicals.

All of the above

Answer

All of the above

This page titled 2.6.5.6: Non-Governmental Regulatory Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **5.6: Non-Governmental Regulatory Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.5.7: Section 5 Final Evaluation

1. _____ was created by UN to harmonize the criteria by which chemicals are classified in terms of their hazards.

Federal Insecticide, Fungicide and Rodenticide Act (FIFRA).

Consumer Product Safety Commission (CPSC).

Globally Harmonized System of Classification and Labelling of Chemicals (GHS).

The US Food and Drug Administration (FDA).

Answer

Globally Harmonized System of Classification and Labelling of Chemicals (GHS).

2. Regulatory agencies generally have specific focus areas that they address, examples include the following EXCEPT:

Food Quality Protection Act (FQPA) addresses the quality of food and chemicals used in an environment.

Consumer Product Safety Commission (CPSC) addresses chemical hazards in consumer products.

Occupational Safety and Health Administration (OSHA) addresses hazardous chemicals in the work place.

Food and Drug Administration (FDA) regulates pharmaceutical drugs used in humans and animals.

Answer

Food Quality Protection Act (FQPA) addresses the quality of food and chemicals used in an environment.

3. Food and Drug Administration (FDA) is a national regulatory agency that serves general population and medical patients while Consumer Product Safety Commission (CPSC) serves general population and environment.

True

False

Answer

False

4. The purpose of Proposition 65 is to enable consumers make informed decisions regarding chemical exposures while its scope is to address chemical exposures to citizens of California that may occur through consumer products, workplace and environmental exposures.

True

False

Answer

True

5. Registration, Evaluation, Authorization and restriction of CHemicals (REACH) is an example of:

Global regulatory toxicology.

Regional regulatory toxicology.

National regulatory toxicology.

State regulatory toxicology.

Answer

Regional regulatory toxicology.

6. One of the advantages of a Non-Governmental Organization (NGO) over government agencies is:

NGOs often conduct regulatory toxicology on a much faster timeframe than government agencies.

NGOs are transparent in terms of decision making while government agencies are not transparent.

NGOs are not only limited to advocacy group but it influences global regulations.

NGOs are officially part of government agency that influence regulatory toxicology.

Answer

NGOs often conduct regulatory toxicology on a much faster timeframe than government agencies.

7. The following are examples of national toxicology regulatory agencies in the U.S. EXCEPT:

Consumer Product Safety Commission (CPSC).

Environmental Protection Agency (EPA).

Texas Commission on Environmental Quality (TCEQ).

Food and Drug Administration (FDA).

Answer

Texas Commission on Environmental Quality (TCEQ).

8. State of Washington Department of Ecology is an agency that sets state toxicology regulations for:

Proposition 65.

Wellhead Protection Program.

Children's Safe Products Act (CSPA).

Risk reduction Program.

Answer

Children's Safe Products Act (CSPA).

9.Global regulatory toxicology brings together member countries to confront common challenges on a global scale, an example is:

United Nations (UN).

Environmental Protection Agency (EPA).

Consumer Product Safety Commission (CPSC).

Occupational Safety and Health Administration (OSHA).

Answer

United Nations (UN).

10. Which of the following statements is true?

CPSC mainly regulates substance in consumer products e.g. shampoos, clothing etc.

EPA mainly regulates substance in occupational environment e.g. air in a factory.

CPSC mainly regulates pharmaceutical drugs and medical devices.

OSHA mainly regulates toxic substances in consumer products.

Answer

CPSC mainly regulates substance in consumer products e.g. shampoos, clothing etc.

11. Which of the following statements is true?

Policies are principles that clarify regulation and carry the force of law.

Guidelines are approaches that interpret regulations and carry the force of law.

Regulations are rules issued by governmental authority and carry the force of law.

Regulations are orders issued by government authority and do not carry the force of law.

Answer

Regulations are rules issued by governmental authority and carry the force of law.

12. Key hazard information to be obtained for existing chemicals to remain in commerce or for new chemicals to enter into commerce falls into three broad hazard categories:

Physical-chemical, animal health and environmental health.

Animal health, human health and environmental health.

Environmental health, physical-chemical and biosafety.

Physical-chemical, human health and environmental health

Answer

Physical-chemical, human health and evironmental health

This page titled 2.6.5.7: Section 5 Final Evaluation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **Section 5 Final Evaluation** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6: Principles of Toxicology

Introduction

ToxTutor is a self-paced tutorial covering key principles of toxicology and was adopted from the National Library of Medicine (NLM) chemical and toxicology databases. While a knowledge of anatomy and physiology is not required for viewing ToxTutor, the **Introduction to the Human Body** from the National Cancer Institute provides a good introduction to the topic.

Topics Covered in this Course

ToxTutor is divided into the following sections:

Each section of ToxTutor contains one or more related content pages. *Next* and *Back* buttons are provided to allow you to navigate through these pages. For more information, see the "Getting Around" section below.
The basic principles of toxicology described in ToxTutor are similar to those taught in university programs and are well described in toxicology literature. A list of the textbooks used as the primary resources for the tutorials is found in the **Bibliography**.

Using ToxTutor

Getting Around

It will take approximately three hours to complete this self-paced tutorial.
There are a variety of ways you can navigate ToxTutor. You can:

- Use the **Glossary** button in the upper right corner of each page to access the **glossary**.
- Click the links in the list above (or any underlined words in any module which are linked) to access sections directly.
- Click the Previous and *Next* links at the bottom of each page to move through the material.
- Use a combination of the above methods to explore the course contents.

Progress

At the top of each page in ToxTutor are links indicating where the current page falls within the overall ToxTutor program. You can click these links to return to this homepage or to the section that contains the page.

Links

Throughout the course, you will encounter links. Any link to a resource outside of ToxTutor, will typically open in a new tab or window. All other links are to other areas within ToxTutor. Links are bold and underlined as seen **here** and **here**.
In addition, clicking on some images may open another external website for more information.

Credits

ToxTutor was adopted from the **U.S National Library of Medicine** in 2021. More Information can be seen in the ToxTutor **Bibliography**.

This page titled 2.6.6: Principles of Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

SECTION OVERVIEW

2.6.6.1: Introduction to Toxicology

Learning Objectives

After completing this lesson, you will be able to:

- Define toxicology and identify adverse effects.
- Recognize the history of toxicology.
- Explain how dose determines whether a substance is a remedy or a poison.
- Differentiate between toxic agents and toxic substances.

In this section...

Topics include:

2.6.6.1.1: What is toxicology?

2.6.6.1.2: Basic Terminology

What We've Covered

In this section, we covered several important concepts:

- Toxicology is the study of adverse effects of chemicals and physical agents on living organisms.
- A xenobiotic is a foreign substance taken into the body.
- A toxic agent is any chemical, physical, or biological agent that can produce an adverse biological effect.
- Toxic substances can be systemic toxicants, which affect the entire body or multiple organs, or organ toxicants, which affect a specific organ or tissues.
- The dose of a substance is the most important determinant of toxicity.

Coming Up...

In the next section, we will explore the concept of dose and its importance to toxicology in greater detail.

This page titled 2.6.6.1: Introduction to Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.6.1.1: What is toxicology?

What is Toxicology?

Toxicology is traditionally defined as "the science of poisons." Over time, our understanding of how various agents can cause harm to humans and other organisms has increased, resulting in a more descriptive definition of toxicologyas "the *study of the adverse effects of chemical, physical, or biological agents on living organisms and the ecosystem, including the prevention and amelioration of such adverse effects."*



Figure 2.6.6.1.1.1: Photos of a Gila monster anda Bothrops jararaca snake. (CC BY-SA 4.0; Leandro Avelar via source)

These adverse effects can take many forms, ranging from immediate death to subtle changes not appreciated until months or years later. They may occur at various levels within the body, such as an organ, a type of cell, or a specific biochemical. Our understanding of how toxic agents damage the body has progressed along with medical knowledge. We now know that various observable changes in anatomic or bodily functions actually result from previously unrecognized changes in specific biochemicals in the body.

Did you know?

The study of toxicology may appear to focus only on poisonings or disasters, but some toxic chemicals can have positive effects. Animal venoms, whether from bees, wasps, snakes, or Gila monsters, are composed of hundreds of chemicals that are being studied as treatments for human diseases.

For example, exantide, a drug derived from Gila monster saliva, has been approved for use in Type 2 diabetes. Captopril, which is used to treat hypertension and heart failure, was developed from studies on the chemical bradykinin-potentiating factor (BPF) in the venom of a South American snake Bothrops jararaca. Melitten, which comes from honeybee venom, is being investigated for its anticancer and antifungal properties.

Figure 2.6.6.1.1.1: Figure 1. Gila monster (top); Bothrops jararaca (bottom)
*(Image Source: Wikipedia, adapted under GNU Free Documentation License - **Original Images**)*

History of Toxicology

Prehistory

Poisonous plants and animals were recognized and their extracts used for hunting or in warfare.

1500 BC

Written records indicate that hemlock, opium, arrow poisons, and certain metals were used to poison enemies or for state executions.

c. 1198

With time, people began to make the connection between exposure to a specific substance and illness or death.

In 1198, Moses Maimonides wrote what may be the first collection of writings on toxicology, *The Treatise on Poisons and Their Antidotes*.

Renaissance and Age of Enlightenment

Certain fundamental toxicology concepts began to take shape. Noteworthy studies include those by Paracelsus in the 16th century and Orfila in the 19th century.

Paracelsus (16th Century)

Determined that specific chemicals were actually responsible for the toxicity of a plant or animal poison.

Documented that the body's response to those chemicals depended on the dose received.

Studies revealed that small doses of a substance might be harmless or beneficial, whereas larger doses could be toxic. This is now known as the **dose-response relationship**, a major concept in toxicology.

"All substances are poisons; there is none which is not a poison. The right dose differentiates a poison and a remedy."
- Paracelsus

Orfila, the founder of toxicology (19th Century)

A Spanish physician, Orfila is often referred to as the **founder of toxicology**.

Orfila was the first to describe a systematic correlation between the chemical and biological properties of poisons of the time.

Orfila demonstrated the effects of poisons on specific organs by analyzing autopsy materials for poisons and tissue damage associated with them.

20th and 21st Centuries

Marked by great advancements in the level of understanding of toxicology.

DNA and various biochemicals that maintain body functions have been discovered.

Our level of knowledge of toxic effects on organs and cells has expanded to the molecular level.

Virtually all toxic effects are recognized as being caused by changes in specific cellular molecules and biochemicals.

Remedy or Poison?

Xenobiotic is the general term that is used for a ***foreign***substance taken into the body. It is derived from the Greek term *xeno* which means "foreigner." Xenobiotics may produce beneficial effects (such as pharmaceuticals) or they may be toxic (such as lead).

As Paracelsus proposed centuries ago, dose differentiates whether a substance will be a remedy or a poison. A xenobiotic in small amounts may be nontoxic and even beneficial, but when the dose is increased, toxic and lethal effects may result.

The following image provides some examples that illustrate this concept.

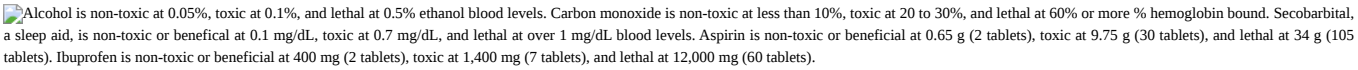


Figure 2.6.6.1.1.2: Figure 3. Examples of varying *doses of the same substance as non-toxic or beneficial, toxic, and lethal (Image Source: Adapted from T. Gossel and J. Bricker, eds)*

This page titled 2.6.6.1.1: What is toxicology? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **1.1: What is toxicology?** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.1.2: Basic Terminology

Toxicology Defined

Toxicology is an evolving medical science and toxicology terminology is evolving with it. Most terms are very specific and will be defined as they appear in the tutorial. However, some terms are more general and used throughout the various sections. The most commonly used terms are introduced in this section.

- **Toxicology** is the study of the adverse effects of chemicals or physical agents on living organisms.
- A **toxicologist** is a scientist who determines the harmful effects of agents and the cellular, biochemical, and molecular mechanisms responsible for the effects.
- **Toxinology**, a specialized area of study, looks at microbial, plant and animal venoms, poisons, and toxins.

Terminology and definitions for materials that cause toxic effects are not always consistently used in the literature. The most common terms are toxicant, toxin, poison, toxic agent, toxic substance, and toxic chemical.

Toxicant, toxin, and poison are often used interchangeably in the literature but there are subtle differences as shown below:

Toxicants:

- Substances producing adverse biological effects of any kind.
- May be chemical or physical in nature.
- Effects may be acute or chronic.

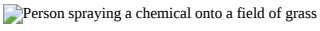


Figure 2.6.6.1.2.1: *Pesticide chemicals are toxicants*
(Image Source: iStock Photos, ©)

Toxins:

- Peptides or proteins produced by living organisms.
- Venoms are toxins injected by a bite or sting.

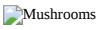


Figure 2.6.6.1.2.2: Amanita muscaria mushroom contains a *neurotoxin*
(Image Source: iStock Photos, ©)

Poisons:

- Toxins produced by organisms.

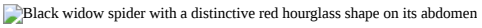


Figure 2.6.6.1.2.3: Black Widow spiders produce a *poison that is a toxin*
(Image Source: Texas Parks & Wildlife, ©)

A **toxic agent** is anything that can produce an adverse biological effect. It may be chemical, physical, or biological in form. For example, toxic agents may be:

Chemical (such as cyanide)

Physical (such as radiation)

Biological (such as snake venom)

The toxicity of the agent is dependent on the dose.

A distinction is made for diseases people get from *living* organisms. Organisms that invade and multiply within another organism and produce their effects by biological activity are not classified as toxic agents but as **biological agents**. An example of this is a virus that damages cell membranes resulting in cell death.

If the invading organisms excrete chemicals which are the basis for their toxicity, the excreted substances are known as **biological toxins**. In that case, the organisms are called **toxic organisms**. A specific example is tetanus. Tetanus is caused by a bacterium, *Clostridium tetani*. The bacteria *C. tetani* itself does not cause disease by invading and destroying cells. Rather, a toxin (neurotoxin) that the bacteria excrete travels to the nervous system and produces the disease (Figure 8).

Toxic Substances

A **toxic substance** is simply a material that has toxic properties. It may be a discrete toxic chemical or a mixture of toxic chemicals. For example, lead chromate, asbestos, and gasoline are all toxic substances. More specifically:

- Lead chromate is a discrete **toxic chemical**.
- Asbestos is a **toxic material** that does not have an exact chemical composition but comprises a variety of fibers and minerals.
- Gasoline s a **toxic substance** rather than a toxic chemical in that it contains a mixture of many chemicals. Toxic substances may not always have a constant composition. The composition of gasoline varies with octane level, manufacturer, time of season, and other factors.

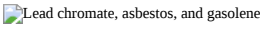


Figure 2.6.6.1.2.4: Examples of *toxic substances: lead chromate (left), asbestos (center), and gasoline (right)*
(Image Source: iStock Photos, ©)

Systemic Toxicants and Organ Toxicants

Toxic substances may be **systemic toxicants** or **organ toxicants**.

A **systemic toxicant** affects the entire body or many organs rather than a specific site. For example, potassium cyanide is a systemic toxicant in that it affects virtually every cell and organ in the body by interfering with the cells' ability to use oxygen.

Toxicants may also affect only specific tissues or organs while not producing damage to the body as a whole. These specific sites are known as the **target organs**or **target tissues**.

- Benzene is a **specific organ toxicant**in that it is primarily toxic to the blood-forming tissues.
- Lead is also a specific organ toxicant; however, it has three **target organs:** the central nervous system, the kidneys, and the hematopoietic system.

A toxicant may affect a specific type of tissue (such as connective tissue) that is present in several organs. The toxic site is then considered the**target tissue**.

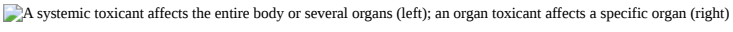


Figure 2.6.6.1.2.5: Systemic toxicant *and organ toxicant*
(Image Source: iStock Photos, ©)

Types of Cells

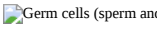
The body is composed of many types of cells, which can be classified in several ways. Table 1 shows examples of one classification of one type of cells.

Cell Types	Examples
Basic structure	cuboidal cells
Tissue type	hepatocytes of the liver
Germ cells	ova and sperm
Somatic cells	non-reproductive cells of the body

Germ cellsare involved in reproduction and can give rise to a new organism. They have only a single set of chromosomes peculiar to a specific sex. Male germ cells give rise to sperm and female germ cells develop into ova. Toxicity to germ cells can cause effects in a developing fetus that lead to outcomes such as birth defects or miscarriage.

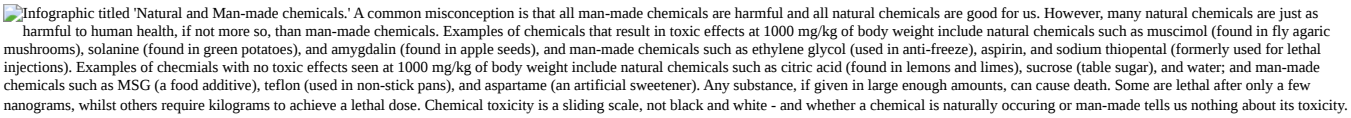
Somatic cellsare all body cells except the reproductive germ cells. (Somatic cells include the "basic structure" and "tissue type" cells listed in Table 1). They have two sets (or pairs) of chromosomes. In an exposed individual, toxicity to somatic cells causes a variety of toxic effects, such as dermatitis, death, and cancer.

Figure 6 illustrates the differences between germ cells and somatic cells.


Germ cells (sperm and egg cells) have a single set of chromosomes. Somatic cells (all other body cells) have two sets, or pairs, of chromosomes.Figure 2.6.6.1.2.6: Germ cells and somatic cells (Image Source: National Human Genome Research Institute, **http://www.genome.gov**)

Natural and Man-Made Chemicals

Often, people mistakenly assume that all man-made chemicals are harmful and natural chemicals are beneficial. In reality, natural chemicals can be just as harmful to human health as man-made chemicals, and in many cases, more harmful. Figure 12 compares the toxicity of several natural and man-made chemicals.


Infographic titled 'Natural and Man-made chemicals.' A common misconception is that all man-made chemicals are harmful and all natural chemicals are good for us. However, many natural chemicals are just as harmful to human health, if not more so, than man-made chemicals. Examples of chemicals that result in toxic effects at 1000 mg/kg of body weight include natural chemicals such as muscimol (found in fly agaric mushrooms), solanine (found in green potatoes), and amygdalin (found in apple seeds), and man-made chemicals such as ethylene glycol (used in anti-freeze), aspirin, and sodium thiopental (formerly used for lethal injections). Examples of checmials with no toxic effects seen at 1000 mg/kg of body weight include natural chemicals such as citric acid (found in lemons and limes), sucrose (table sugar), and water; and man-made chemicals such as MSG (a food additive), teflon (used in non-stick pans), and aspartame (an artificial sweetener). Any substance, if given in large enough amounts, can cause death. Some are lethal after only a few nanograms, whilst others require kilograms to achieve a lethal dose. Chemical toxicity is a sliding scale, not black and white - and whether a chemical is naturally occuring or man-made tells us nothing about its toxicity.
Figure 2.6.6.1.2.7: Natural and man-made *chemicals (Image Source: **http://pbs.twimg.com/media/CTn6HlDWwAEB1Jc.png:large**, Creative Commons license)*

This page titled 2.6.6.1.2: Basic Terminology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **1.2: Basic Terminology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.2: Dose and Dose Response

Learning Objectives

After completing this lesson, you will be able to:

- Define *dose* and explain its importance in determining toxicity.
- Analyze and compare dose-response curves.
- Describe the safety or toxicity of substances based on specific dose levels.

In this section...

Topics include:

2.6.6.2.1: Dose and It's Impact on Toxicity

2.6.6.2.2: The Dose Response Relationship

2.6.6.2.3: Dose Estimates of Toxic Effects

2.6.6.2.4: Determining the Safety of a Drug

2.6.6.2.5: NOAEL and LOAEL

What We've Covered

In this section, we explored the following main points:

- Dose is the amount of a substance administered; however, several parameters are required to characterize exposure to xenobiotics, including the:
  - Number of doses
  - Frequency of doses
  - Total time period of exposure
- The dose-response relationship helps establish causality, or that the chemical induced the observed effects; the threshold effect, or the lowest dose that induced effects; and the slope, or the rate at which effects increase with dose increases.
- Estimating doses for toxic effects involves:
  - Lethal Doses/Concentrations, such as LD0, LD10, and LC50, which denote doses or concentrations that are expected to lead to death in specific percentages of a population.
  - Effective Doses, such as ED50 and ED90, which denote doses that are effective in achieving a desired endpoint in specific percentages of a population.
  - Toxic Doses, such as TD0 and TD50, which denote doses that cause adverse toxic effects in specific percentages of a population.
- The Therapeutic Index (TI) compares the effective dose to the toxic dose of a drug.
- The Margin of Safety (MOS) compares the toxic dose to 1% of the population to the effective dose to 99% of the population.
- NOAEL is the **highest** dose at which there **is no** observed toxic effect.
- LOAEL is the **lowest** dose at which there **is** an observed toxic effect.

Coming Up...

In the next section, we will take a closer look at various types of toxic effects.

This page titled 2.6.6.2: Dose and Dose Response is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.6.2.1: Dose and It's Impact on Toxicity

Dose Defined

Dose by definition is the amount of a substance administered at one time. However, other parameters are needed to characterize the exposure to xenobiotics. The most important are the number of doses, frequency, and total time period of the treatment.

For example:

- 650 -3 g)." tabindex="0">mg acetaminophen (Tylenol® products) as a single dose.
- 500 -3 g)." tabindex="0">mg penicillin every 8 hours for 10 days.
- 10 -3 g)." tabindex="0">mg DDT per day for 90 days.

Substances can enter the body from either:

1. Encountering them in the environment (**exposure**).
2. Intentionally consuming or administering a certain quantity of a substance.

Environments in which xenobiotics are present include outdoor air, indoor air, and water. Xenobiotics can travel into the body through the skin, eyes, lungs, and digestive tract. **Exposure to a xenobiotic** can occur in any environment where a substance can enter the:

- Skin through dermal absorption (air and water).
- Respiratory tract through inhalation.
- Digestive tract through ingestion.



Figure 2.6.6.2.1.1: Dermal
(Image Source: ORAU, ©)



Figure 2.6.6.2.1.2: *Inhalation*
(Image Source: ORAU, ©)



Figure 2.6.6.2.1.3: *Ingestion*
(Image Source: ORAU, ©)

A dose can be considered either:

1. A measurement of environmental exposures.
2. The amount of a substance administered over a period of time.

Types of doses include:

- **Absorbed dose** — the amount of a substance that entered the body through the skin, eyes, lungs, or digestive tract and was taken up by organs or particular tissues. Absorbed dose can also be called **internal dose**.
- **Administered dose** — the quantity administered usually orally or by injection (note that an administered dose taken orally may not necessarily be absorbed).
- **Total dose** — the sum of all individual doses.

Not all substances that enter the body are necessarily absorbed by it. This concept applies to water intake. When a person drinks a large quantity of water at one time, some of it is absorbed while the rest of the water is eliminated.

If an individual drinks 1 liter of water every hour for 3 hours, each *administered dose* would be 1 liter. The *total dose* would be the amount the person drank over the time period that the water was consumed. The *absorbed dose, however,* would likely be less than the total dose because it would depend on how much of the water the individual's body absorbed which can be affected by various factors. The water intake example is represented in Table 1.

Doses	Administered Dose	Total Dose	Absorbed Dose
Dose 1 (8:00 a.m.): 1 L water	1 L	1 L	Less than 1 L
Dose 2 (9:00 a.m.): 1 L water	1 L	2 L	Less than 2 L
Dose 3 (10:00 a.m.): 1 L water	1 L	3 L	Less than 3 L

The terms for types of doses help account for the amount of a substance that entered the body by different means, but the amount *absorbed* is what is most important.

In a later section, we will review specifics about how the body handles substances after they enter the body.

Fractionating Doses

Fractionating a total dose usually decreases the probability that the total dose will cause toxicity. The reason is that the body often can repair the effect of each**subtoxic** dose if sufficient time elapses before the next dose is received. In that case, a total dose that would be harmful if received all at once is non-toxic when administered over a period of time. For example, 30 -3 g)." tabindex="0">mg of strychnine swallowed at one time could be fatal to an adult whereas 3 -3 g)." tabindex="0">mg of strychnine swallowed each day for 10 days is not considered a fatal dose.

The **units** used in toxicology are basically the same as those used in medicine. The **gram (g)** is the standard unit. Because most exposures are in smaller quantities, the **-3 g)." tabindex="0">milligram (-3 g)." tabindex="0">mg)** is commonly used. For example, the common adult dose of acetaminophen is 650 -3 g)." tabindex="0">mg.

Importance of Age, Body Size, and Time

A person's age and body size affect the clinical and toxic effects of a given dose. Age and body size usually are connected, particularly in children. This relationship is important because a person's body size can affect the burden that a substance has on it. For example, a 650--3 g)." tabindex="0">mg dose of acetaminophen is typical for adults but it would be toxic to young children. Therefore, a tablet of an acetaminophen product designed for children (Children's Tylenol®) contains only 80 -3 g)." tabindex="0">mg of the drug.

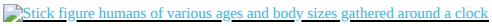


Figure 2.6.6.2.1.4: Age, body size, and time are key factors when considering the clinical and toxic effects of a dose
(Image Source: iStock Photos, ©)

One way to compare the effectiveness of a dose and its toxicity is to assess the amount of a substance administered with respect to **body weight**. A common dose measurement is **mg/kg** which stands for -3 g)." tabindex="0">mg of substance per kg of body weight. Another method used to compare doses among different species is to use **body surface area**, rather than simply body weight.

Units

Because some xenobiotics are toxic in quantities much smaller than the -3 g)." tabindex="0">milligram, smaller fractions of the gram, such as **-6) of a gram." tabindex="0">microgram (µg)**. Table 2 shows other units.

Unit	Gram Equivalents	Exp. Form
Kilogram (kg)	1000.0 g	10^3g
Gram (g)	1.0 g	1 g
-3 g)." tabindex="0">Milligram (-3 g)." tabindex="0">mg)	0.001 g	10^{-3} g
-6) of a gram." tabindex="0">Microgram (µg)	0.000,001 g	10^{-6} g
-9 g)." tabindex="0">Nanogram (-9 g)." tabindex="0">ng)	0.000,000,001 g	10^{-9} g
-12 g)." tabindex="0">Picogram (-12 g)." tabindex="0">pg)	0.000,000,000,001 g	10^{-12} g
Femtogram (fg)	0.000,000,000,000,001 g	10^{-15} g

Table 2.6.6.2.1.1: *Various units, their gram equivalents, and their exponential form*

Concentration

Environmental exposure units are expressed as the amount of a xenobiotic in a unit of the media, which could be liquid, solid, or air. **Concentration** is the amount of a substance found in a certain amount of another substance, such as water, air, soil, food, blood, hair, urine, or breath. For example, the weight of a toxic substance found in a certain weight of food is indicated as a measure of concentration rather than the total amount. Knowing how concentrated the toxic substance is in a sample of food that weighs 100 g allows for easy comparison when testing for that toxic substance in other samples of food that weigh more or less than 100 g.

Figure 7 illustrates this concept. The two glasses contain samples of juice that are being tested for contamination with lead. The volume of juice in Glass A is 100 mL and the volume of juice in Glass B is 50 mL. The concentration of lead is the same in both samples of juice: 20 parts per billion (ppb). The total amount of lead would be higher in Glass A but the concentration of lead per unit volume is the same in both glasses.



Figure 2.6.6.2.1.5: The concentration of lead is the same in samples with different volumes, but the total amount of lead in each is not
(Image Source: Adapted from iStock Photos, ©)

Assessing Exposure

An individual's exposure to a substance can be assessed based on the relationship between the person's *body weight* and these factors:

- *Concentration* of the substance in the environmental media (for example, in µg/ml).
- Amount of the substance taken into the body.
- Duration and frequency of individual events during which the body was in contact with the environmental media.

Environmental exposure units used in toxicology include:

- -3 g)." tabindex="0">mg/liter (**-3 g)." tabindex="0">mg/L**) for liquids.
- -3 g)." tabindex="0">mg/gram (**-3 g)." tabindex="0">mg/g**) for solids.
- -3 g)." tabindex="0">mg/cubic meter (**-3 g)." tabindex="0">mg/m^3**) for air.

Smaller units are used as needed; for example, **-6) of a gram." tabindex="0">µg/mL**. Other commonly used dose units for substances in media are parts per million **(ppm)**, parts per billion **(ppb)**, and parts per trillion **(ppt)**. When smaller units are used to

quantify exposure, the mg/kg/day unit can be adapted to the smaller unit. For example, parts per billion per kg per day (ppb/kg/day) could be used.

An important thing to remember is that the use of a small dose unit is *not* related to the burden a substance has on the body. An exposure unit describes only the *quantity* of the substance.

This page titled 2.6.6.2.1: Dose and It's Impact on Toxicity is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **2.1: Dose and It's Impact on Toxicity** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.2.2: The Dose Response Relationship

The Dose-Response Relationship

The dose-response relationship is an essential concept in toxicology. It correlates exposures with changes in body functions or health.

In general, the higher the dose, the more severe the response. The dose-response relationship is based on observed data from experimental animal, human clinical, or cell studies.

Knowledge of the dose-response relationship establishes:

- **Causality** — that the chemical has induced the observed effects.
- The **threshold effect** — the lowest dose where an induced effect occurs.
- The **slope** for the dose response — the *rate* at which injury builds up.

Within a population, the majority of responses to a toxicant are similar; however, there are differences in how responses may be encountered – some individuals are susceptible and others resistant. As demonstrated in Animation 1, a graph of the individual responses can be depicted as a bell-shaped standard distribution curve. There is a wide variance in responses as demonstrated by the mild reaction in resistant individuals, the typical response in the majority of individuals, and the severe reaction in sensitive individuals.

*Animation 1: A graph of individual responses to a substance, which generally take the form of a bell-shaped curve (**view full-text, PDF version**)*

The **dose-response curve** is a visual representation of the response rates of a population to a range of doses of a substance, as demonstrated in Animation 2 (available at ToxTutor). *The graph of a dose-response relationship typically has an "s" shape. (**view full-text, PDF version**)*

Knowledge Check

1. The quantity of a substance administered to an individual over a period of time or in several individual doses is known as the:

- ◯ administered dose
- ◯ absorbed dose
- ◯ total dose

Answer

total dose

It is the quantity of a substance administered to an individual over a period of time or in several individual doses. It is particularly important when evaluating cumulative poisons.

2. Fractionation of a total dose so that the total amount administered is given over a period of time usually results in:

- ◯ decreased toxicity
- ◯ increased toxicity

Answer

decreased

toxicity

Fractionation of a total dose so that the total amount administered is given over a period of time usually results in **decreased toxicity**. This applies to most forms of toxicity but not necessarily to carcinogenicity or mutagenicity.

3. The usual dosage unit that incorporates the amount of material administered or absorbed in accordance with the size of the individual over a period of time is:

- ◯ PPM/hour
- ◯ mg/kg/day
- ◯ kg/100 lb/week

Answer

mg/kg/day

The usual dosage unit that incorporates the amount of material administered or absorbed in accordance with the size of the individual over a period of time is **mg/kg/day**. In some cases, much smaller dosage units, such as µg/kg/day, are used.

4. The dose at which a toxic effect is first encountered is called the:

- ◯ median toxic dose
- ◯ absorbed dose
- ◯ threshold dose

Answer

threshold dose

5. The dose-response relationship helps a toxicologist determine:

- ◯ whether exposure has caused an effect
- ◯ the threshold dose
- ◯ the rate of increasing effect with increasing dose levels
- ◯ all of the above

Answer

all of the above

The dose-response relationship demonstrates whether any effect has occurred, the threshold dose, and the rate at which the effect increases with increasing dose levels.

This page titled 2.6.6.2.2: The Dose Response Relationship is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **2.2: The Dose Response Relationship** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.2.3: Dose Estimates of Toxic Effects

Dose Estimates

Dose-response curves are used to derive dose estimates of chemical substances.

Historically, **LD50** (Lethal Dose 50%) has been a common dose estimate for acute toxicity. It is a **statistically derived maximum dose** at which 50% of the group of organisms (rat, mouse, or other species) would be expected to die. LD50 testing is no longer the recommended method for assessing toxicity because of the ethics of using large numbers of animals, the variability of responses in animals and humans, and the use of mortality as the only endpoint. Regulatory agencies use LD50 only if it is justified by scientific necessity and ethical considerations.

The Three Rs

The current practice for estimating acute toxicity emphasizes the following approaches, known as the Three Rs:

1. **Replacing** animals in science by *in vitro*, ***in silico***, and other approaches.
2. **Reducing** the number of animals used. For example, the oral LD50 approach has been replaced in some circumstances by an **up-and-down method** in which animals are dosed one at a time.
3. **Refining** care and procedures to minimize pain and distress.

Other dose estimates also may be used.

Lethal Doses/Concentrations

- **Lethal Dose 0% (LD0)** — represents the dose at which no individuals are expected to die. This is just below the threshold for lethality.
- **Lethal Dose 10% (LD10)** — refers to the dose at which 10% of the individuals will die.
- **Lethal Concentration 50% (LC50)** — for inhalation toxicity, air concentrations are used for exposure values. The LC50 refers to the calculated concentration of a gas lethal to 50% of a group. Occasionally LC0 and LC10 are also used.

Effective Doses (EDs)

Effective Doses(**EDs**) are used to indicate the effectiveness of a substance. Normally, effective dose refers to a beneficial effect such as relief of pain. It may also stand for a harmful effect such as paralysis. Thus, the specific endpoint must be indicated. The usual terms are:

Term	Effective for this percentage of the population
ED0	0%
ED10	10%
ED50	50%
ED90	90%

Table 2.6.6.2.3.1. Typical terms for describing effective doses

Toxic Doses (TDs)

Toxic Doses (**TDs**) are used to indicate doses that cause adverse toxic effects. The usual dose estimates include:

Term	Toxic to this percentage of the population
TD0	0%
TD10	10%
TD50	50%
TD90	90%

Table 2.6.6.2.3.2. Typical terms for describing toxic doses

Determining the Relative Safety of Pharmaceuticals

Toxicologists, pharmacologists, and others use **effective**and **toxic dose**levels to determine the relative safety of pharmaceuticals. As shown in Figure 1, two dose-response curves are presented for the same drug, one for effectiveness and the other for toxicity. In this case, a dose that is 50% to 75% effective does not cause toxicity. However, a 90% effective dose may result in a small amount of toxicity.


Graph of two dose-response curves for the same drug - one for effectiveness and one for toxicity. ED10, ED50, ED90, TD10, TD50, TD90 are marked on the graph. A 50-75% effective dose does not cause toxicity in this example, but a 90% effective dose could result in a small amount of toxicity.

Figure 2.6.6.2.3.1. Dose-response curves representing effective dose and toxic dose for the same drug
(Image Source: NLM)

It should be noted that a desired effect in a drug is often an undesired effect with an environmental chemical.

This page titled 2.6.6.2.3: Dose Estimates of Toxic Effects is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **2.3: Dose Estimates of Toxic Effects** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.2.4: Determining the Safety of a Drug

Do you know?

What are measures for describing the safety of a drug? This section describes the:

- Therapeutic index
- Margin of safety

Figure 2.6.6.2.4.1: (Image Source: iStock Photos, ©)

Therapeutic Index

The **Therapeutic Index (TI)** is used to compare the therapeutically effective dose to the toxic dose of a pharmaceutical agent. The TI is a statement of relative safety of a drug. It is the ratio of the dose that produces toxicity to the dose needed to produce the desired therapeutic response. The common method used to derive the TI is to use the 50% dose-response points, including TD50 (toxic dose) and ED50 (effective dose).

For example, if the TD50 is 200 and the ED50 is 20 -3 g)." tabindex="0">mg, the TI would be 10.

Figure 1).

Figure 2.6.6.2.4.2. A higher value on the Therapeutic Index indicates a more favorable safety profile
(Image Source: ORAU, ©)

However, the use of the **ED50**and **TD50**doses to derive the **TI**may be misleading about a drug's safety, depending on the slope of the dose-response curves for therapeutic and toxic effects. To overcome this deficiency, toxicologists often use another term to denote the safety of a drug: the Margin of Safety.

Margin of Safety (MOS)

The **Margin of Safety (MOS)** is usually calculated as the ratio of the toxic dose to 1% of the population (TD01) to the dose that is 99% effective to the population (ED99).

Figure 2 shows the relationship between effective dose response and toxic dose response. The shaded area represents the doses at which the substance produces an effective dose response while the toxic dose response remains below the TD50. The slope of a curve shows how dose increases result in responses to the effective or toxic dose.

Figure 2.6.6.2.4.3. Relationship between effective dose response and toxic dose response
(Image Source: NLM)

Because of differences in slopes and threshold doses, low doses may be effective without producing toxicity. Although more patients may benefit from higher doses, that is offset by the probability that toxicity will occur.

The toxicity of various substances can be compared using the slopes for each curve (Figure 3).

Figure 2.6.6.2.4.4. Comparison of the toxicity of two substances
(Image Source: NLM)

For some substances, a small increase in dose causes a large increase in response, which is seen in Toxicant A's steep slope. For other substances, a much larger increase in dose is required to cause the same increase in response, as indicated in Toxicant B's shallow slope.

This page titled 2.6.6.2.4: Determining the Safety of a Drug is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **2.4: Determining the Safety of a Drug** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.2.5: NOAEL and LOAEL

NOAEL and LOAEL

Results from research studies establish the highest doses at which no toxic effects were identified and the lowest doses at which toxic or adverse effects were observed. The terms often used to describe these outcomes are:

- **No Observed Adverse Effect Level (NOAEL)**
- **Lowest Observed Adverse Effect Level (LOAEL)**

These terms refer to the actual doses used in human clinical or experimental animal studies. They are defined as follows:

- NOAEL — *Highest dose* at which there *was not* an observed toxic or adverse effect.
- LOAEL — *Lowest dose* at which there *was* an observed toxic or adverse effect.

Figure 1 shows a dose-response curve where the NOAEL occurs at 10 mg and the LOAEL occurs at 18 mg.



Figure 2.6.6.2.5.1. A dose-response curve showing doses where the NOAEL and LOAEL occur for a substance
(Image Source: NLM)

Sometimes the terms **No Observed Effect Level (NOEL)** and **Lowest Observed Effect Level (LOEL)** are also used. NOELs and LOELs do not necessarily imply toxic or harmful effects and can be used to describe beneficial effects of substances.

The NOAEL, LOAEL, NOEL, and LOEL are commonly used in risk assessments and research. For example, **this U.S. Food and Drug Administration (FDA) publication** for industry describes a process for estimating the maximum safe starting dose of drugs tested in clinical trials. It provides extensive information about these concepts and their utility when developing new drugs.

NOEALs and LOAELs are also included in the **Noncarcinogenic Risk Assessment section** where they are applied using the benchmark dose (BMD) method.

Knowledge Check

1. Which of the following is **not** one of the Three Rs of estimating acute toxicity?

- ◯ Replace animals in science by *in vitro*, *in silico*, and other approaches
- ◯ Reduce the number of animals used
- ◯ Refine care and procedures to minimize pain and distress
- ◯ Randomize the selected test animals

Answer

Randomize the selected test

animals

The Three Rs involve replacing animals in science by *in vitro*, *in silico*, and other approaches; reducing the number of animals used in testing; and refining care and procedures to minimize pain and distress.

2.

The Therapeutic Index (TI) is used to:

- ◯ Compare the therapeutically effective dose to the toxic dose of a pharmaceutical agent
- ◯ Calculate the lethal dose level for different pharmaceuticals
- ◯ Compare the lethal dose to the therapeutically effective dose of a pharmaceutical agent

Answer

Compare the therapeutically effective dose to the toxic dose of a pharmaceutical agent

3. The Margin of Safety (MOS) of a drug is the:

- ◯ Amount of a pharmaceutical that can be given before toxicity first appears
- ◯ Difference between the Effective Dose to 50% of the population (ED50) and the Toxic Dose to 50% of the population (TD50)
- ◯ Ratio of the Toxic Dose to 1% of the population (TD01) to the Effective Dose to 99% of the population (ED99)

Answer

◯ Ratio of the Toxic Dose to 1% of the population (TD01) to the Effective Dose to 99% of the population (ED99)

4. The No Observed Adverse Effect Level (NOAEL) is the:

- ◯ *Lowest dose* at which there *was no* observed toxic or adverse affect
- ◯ *Highest dose* at which there *was no* observed toxic or adverse effect
- ◯ *Highest dose* at which there *was* an observed adverse effect

Answer

Highest dose at which there *was no* observed toxic or adverse effect

5. The Lowest Observed Adverse Effect Level (LOAEL) is the:

- ◯ *Lowest dose* at which there *was* an observed toxic or adverse effect
- ◯ *Lowest dose* at which there *was no* observed toxic or adverse effect
- ◯ *Highest dose* at which there *was* an observed adverse effect

Answer

Lowest dose at which there *was* an observed toxic or adverse effect

This page titled 2.6.6.2.5: NOAEL and LOAEL is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **2.5: NOAEL and LOAEL** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.3: Toxic Effects

Learning Objectives

After completing this lesson, you will be able to:

- Explain factors that influence the toxicity of a substance.
- Define types of systemic and organ-specific toxic effects.

In this section...

Topics include:

2.6.6.3.1: Types of Toxic Effects

2.6.6.3.2: Factors Affecting Toxicity

2.6.6.3.3: Systemic Toxic Effects

2.6.6.3.4: Organ Specific Toxic Effects

Did you know?

In December 1984, the world's worst industrial accident occurred in Bhopal, India. More than 40 tons of **methyl isocyanate** leaked from a pesticide plant, killing thousands of people and injuring hundreds of thousands. Follow-up studies have shown that the incident caused increased mortality and continued effects on health, including airway disease, eye diseases, and pregnancy losses.

The company involved in the leak tried to distance itself from the accident and prevent those affected from learning the true nature of the accident. The legal case went on for years. Eventually, families of the dead received an average of about $2,200. While the company ceased operation at its Bhopal plant after the disaster, it did not clean up the site completely. The plant continues to leak several toxic chemicals and heavy metals into local aquifers.



Figure 1. Survivors of the Bhopal disaster of 1984 protest over the mishandling of the disaster
Bhopal disaster protestors. [Photo]. In Encyclopædia Britannica. Retrieved from **http://www.britannica.com/event/Bhopal-disaster/images-videos/Survivors-of-the-1984-deadly-industrial-accident-in-Bhopal-India/192038**

What We've Covered

This section made the following main points:

- Toxicity can result from adverse cellular, biochemical, or macromolecular changes.
- Some chemicals affect only specific target organs; others can damage any cell or tissue they contact.
- Chemicals can affect organisms by multiple mechanisms and at the molecular level, leading to modern approaches such as Adverse Outcome Pathways (AOPs) and Mechanism of Actions (MOAs).
- Several factors influence toxicity, including form and innate chemical activity, dosage, exposure route, species, life stage, gender, absorption ability, metabolism, distribution, excretion, health and nutritional status, the presence of other chemicals, and circadian rhythms.
- Systemic toxic effects, which can occur at multiple sites, include:
  - Acute toxicity, which occurs almost immediately (seconds/minutes) after a single dose or series of doses within 24 hours.
  - Subchronic toxicity, which results from repeated exposure for several weeks or months.
  - Chronic toxicity, which damages specific organ systems over the course of many months or years.
  - Carcinogenicity, or abnormal cell growth and differentiation that can lead to cancer.

- Developmental toxicity, which adversely affects the developing embryo or fetus.
- Genetic toxicity, caused by damage to DNA and altered genetic expressions.

- Organ specific toxic effects include:
  - Blood/cardiovascular toxicity, affecting the blood, bone marrow, or heart.
  - Dermal toxicity, impacting the skin.
  - Epigenetic alterations, changing genetic programming.
  - Optical toxicity, adversely affecting the eyes.
  - Hepatotoxicity, impacting the liver, bile duct, or gall bladder.
  - Immunotoxicity, affecting the immune system.
  - Nephrotoxicity, affecting the kidneys.
  - Neurotoxicity, impacting the central nervous system.
  - Reproductive toxicity, damaging the reproductive system.
  - Respiratory toxicity, affecting the respiratory system.

Coming Up...

In the next section, we will look at effects that can occur when two or more chemicals interact.

This page titled 2.6.6.3: Toxic Effects is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.6.3.1: Types of Toxic Effects

Types of Toxic Effects

Many factors play a potential role in toxicity. The dosage (or amount of exposure) is the most important factor. A well-known saying, "**the dose makes the poison**" speaks to this principle.

Toxicity can result from adverse cellular, biochemical, or macromolecular changes. Some examples are noted below.

Many chemicals distribute in the body and often affect only specific **target organs**. However, other chemicals can damage any cell or tissue that they contact. The target organs that are affected may vary depending on dosage and route of exposure. For example, the central nervous system may be the target organ for toxicity from a chemical after acute exposure whereas the liver may be affected after chronic exposures.



Figure 2.6.6.3.1.1. Central nervous system (left); Liver (right)
(Image Source: iStock Photos, ©)

Chemicals can cause many types of toxicity by a variety of mechanisms. Some act locally such as when direct exposure triggers skin or eye irritation, whereas other chemical cause systemic effects in the body in sites remote from where the actual exposure occurred. Toxicity can act directly affect subcellular components, such as cell receptors, or it can cause problems at the cellular level, such as with exposures to caustic or corrosive substances.

For example, chemicals might:

- Themselves be toxic or require metabolism (chemical change within the body) before they cause toxicity.
- Cause damage leading to fibrosis as the body attempts to repair the toxicity.
- Damage or disrupt an enzyme system or protein synthesis.
- Produce reactive chemicals in cells.
- Cause changes in hormone signaling or other effects.
- Produce DNA damage or epigenetic changes.

Some chemicals may also act indirectly by:

- Modifying an essential biochemical function.
- Interfering with nutrition.
- Altering a physiological mechanism.



Figure 2.6.6.3.1.2. Chemicals can have a wide range of toxic effects
(Image Source: iStock Photos, ©)

Did you know?

Mercury is a naturally occurring heavy metal. **Methylmercury** , the most common organic mercury compound, can be formed in water and soil by bacteria. It builds up in the tissues of fish. Exposure to high levels of mercury and mercury compounds can cause death or permanently damage the brain and kidneys.

In the late 1950s, people living around Japan's Minamata Bay developed symptoms of severe methylmercury poisoning, some of whom died. Children exposed *in utero* were born with disabilities. Investigations showed that heavily contaminated sludge from a factory had been released into the bay, contaminating fish and shellfish. People who ate the fish and shellfish became ill. The events led to a better understanding of industrial pollution and how heavy metals can accumulate in systems.

In January 2013 the **Minamata Convention on Mercury** global treaty was agreed to by an intergovernmental committee. It seeks to protect human health and the environment from the adverse effects of mercury.

Figure 2.6.6.3.1.3. Fish and shellfish became contaminated with mercury, causing severe methylmercury poisoning in people and animals who consumed them
(Image Source: iStock Photos, ©)

Because chemicals can affect organisms by different mechanisms and at the molecular level, there are new ways to conduct toxicity testing.

An emerging approach is to use **Adverse Outcome Pathways (AOPs)**, which evaluate changes in normal cellular pathways. AOPs reflect the move away from high-dose studies in laboratory animals for toxicity testing to *in vitro* methods that evaluate changes in normal cellular pathways using human-relevant cells or tissues.

Other terms that describe changes resulting from the exposure of a living organism to a substance include mode of action (MoA) and mechanism of action (MOA).

- **Mode of action (MoA)** (older term) — describes a functional or anatomical change at the *cellular* level.
- **Mechanism of action (MOA)** — describes such changes at the *molecular* level.

More information about toxicity testing can be found in the **Hazard Identification** section.

This page titled 2.6.6.3.1: Types of Toxic Effects is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1: Types of Toxic Effects** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.3.2: Factors Affecting Toxicity

Factors Influencing Toxicity

In some instances, individuals can have unpredictable reactions, or **idiosyncratic responses**, to a drug or other substance. An idiosyncratic response is uncommon, and it is sometimes impossible to understand whether it is the result of a genetic predisposition or has some other cause such as the status of the immune system. It could result in an abnormally small or short, or abnormally large or long response to the drug or other substance. Or, the response could be qualitatively different than what has been observed in most other individuals.

The toxicity of a substance usually depends on the following factors:

- Form and innate chemical activity
- Dosage, especially dose-time relationship
- Exposure route
- Species
- Life stage, such as infant, young adult, or elderly adult
- Gender
- Ability to be absorbed
- Metabolism
- Distribution within the body
- Excretion
- Health of the individual, including organ function and pregnancy, which involves physiological changes that could influence toxicity
- Nutritional status
- Presence of other chemicals
- Circadian rhythms (the time of day a drug or other substance is administered)

Factors Related to the Substance

Form and Innate Chemical Activity

The **form** of a substance may have a profound impact on its toxicity especially for metallic elements, also termed heavy metals. For example, the toxicity of mercury vapor differs greatly from methyl mercury. Another example is chromium. Cr^{3+} is relatively nontoxic whereas Cr^{6+} causes skin or nasal corrosion and lung cancer.

The **innate chemical activity** of substances also varies greatly. Some can quickly damage cells causing immediate cell death. Others slowly interfere only with a cell's function. For example:

- Hydrogen cyanide binds to the enzyme cytochrome oxidase resulting in cellular hypoxia and rapid death.
- Nicotine binds to cholinergic receptors in the central nervous system (CNS) altering nerve conduction and inducing gradual onset of paralysis.

Dosage

The **dosage** is the most important and critical factor in determining if a substance will be an *acute* or a *chronic* toxicant. Virtually all chemicals can be acute toxicants if sufficiently large doses are administered. Often the toxic mechanisms and target organs are different for acute and chronic toxicity. Examples are:

Toxicant	Acute Toxicity	Chronic Toxic Effects
Ethanol	CNS depression	Liver cirrhosis
Arsenic	Gastrointestinal damage	Skin/liver cancer

Table 2.6.6.3.2.1. Examples of acute and chronic toxicity

Exposure Route

The way an individual comes in contact with a toxic substance, or **exposure route**, is important in determining toxicity. Some chemicals may be highly toxic by one route but not by others. Two major reasons are differences in absorption and distribution within the body. For example:

- Ingested chemicals, when absorbed from the intestine, distribute first to the liver and may be immediately detoxified.
- Inhaled toxicants immediately enter the general blood circulation and can distribute throughout the body prior to being detoxified by the liver.

Different target organs often are affected by different routes of exposure.



Figure 2.6.6.3.2.1. Ingestion
(Image Source: ORAU, ©)



Figure 2.6.6.3.2.2. Inhalation
(Image Source: ORAU, ©)

Absorption

The **ability to be absorbed** is essential to systemic toxicity. Some chemicals are readily absorbed and others are poorly absorbed. For example, nearly all alcohols are readily absorbed when ingested, whereas there is virtually no absorption for most polymers. The rates and extent of absorption may vary greatly depending on the form of a chemical and the route of exposure to it. For example:

- Ethanol is readily absorbed from the gastrointestinal tract but poorly absorbed through the skin.
- Organic mercury is readily absorbed from the gastrointestinal tract; inorganic lead sulfate is not.

Factors Related to the Organism

Species

Toxic responses can vary substantially depending on the species. Most differences between species are attributable to differences in metabolism. Others may be due to anatomical or physiological differences. For example, rats cannot vomit and expel toxicants before they are absorbed or cause severe irritation, whereas humans and dogs are capable of vomiting.

Selective toxicity refers to species differences in toxicity between two species simultaneously exposed. This is the basis for the effectiveness of pesticides and drugs. For example:

- An insecticide is lethal to insects but relatively nontoxic to animals.
- Antibiotics are selectively toxic to microorganisms while virtually nontoxic to humans.

Life Stage

An individual's age or **life stage** may be important in determining his or her response to toxicants. Some chemicals are more toxic to infants or the elderly than to young adults. For example:

- Parathion is more toxic to young animals.
- Nitrosamines are more carcinogenic to newborn or young animals.



Figure 2.6.6.3.2.3. An individual's life stage can impact that person's response to toxicants
(Image Source: iStock Photos, ©)

Gender

Gender can play a big role in influencing toxicity. Physiologic differences between men and women, including differences in pharmacokinetics and pharmacodynamics, can affect drug activity.

In comparison with men, pharmacokinetics in women generally can be impacted by their lower body weight, slower gastrointestinal motility, reduced intestinal enzymatic activity, and slower kidney function (glomerular filtration rate). Delayed gastric emptying in women may result in a need for them to extend the interval between eating and taking medications that require absorption on an empty stomach. Other physiologic differences between men and women also exist. Slower renal clearance in women, for example, may result in a need for dosage adjustment for drugs such as digoxin that are excreted via the kidneys.

In general, pharmacodynamic differences between women and men include greater sensitivity to and enhanced effectiveness, in women, of some drugs, such as beta blockers, opioids, and some antipsychotics.

Studies in animals also have identified gender-related differences. For example:

- Male rats are 10 times more sensitive than females to liver damage from DDT.
- Female rats are twice as sensitive to parathion as are male rats.



Figure 2.6.6.3.2.4. *Gender symbols for female (left) and male (right)*
(Image Source: iStock Photos, ©)

Metabolism

Metabolism, also known as biotransformation, is the conversion of a chemical from one form to another by a biological organism. Metabolism is a major factor in determining toxicity. The products of metabolism are known as metabolites. There are two types of metabolism:

1. Detoxification
2. Bioactivation

In **detoxification**, a xenobiotic is converted to a less toxic form. This is a natural defense mechanism of the organism. Generally, detoxification converts lipid-soluble compounds to polar compounds.

In **bioactivation**, a xenobiotic may be converted to more reactive or toxic forms. Cytochrome P-450 (CYP450) is an example of an enzyme pathway used to metabolize drugs. In the elderly, CYP450 metabolism of drugs such as phenytoin and carbamazepine may be decreased. Therefore, the effect of those drugs may be less pronounced. CYP450 metabolism also can be inhibited by many drugs. Risk of toxicity may be increased if a CYP450 enzyme-inhibiting drug is given with one that depends on that pathway for metabolism.

There is awareness that the gut microbiota can impact the toxicity of drugs and other chemicals. For example, gut microbes can metabolize some environmental chemicals and bacteria-dependent metabolism of some chemicals can modulate their toxicity. Also, environmental chemicals can alter the composition and/or the metabolic activity of the gastrointestinal bacteria, thus contributing in a meaningful way to shape an individual's microbiome. The study of the consequences of these changes is an emerging area of toxicology.

Learn more about human exposure to pollutants and their interaction with the GI microbiota.

Learn more about the microbiome and toxicology.

Distribution Within the Body

The **distribution** of toxicants and toxic metabolites throughout the body ultimately determines the sites where toxicity occurs. A major determinant of whether a toxicant will damage cells is its lipid solubility. If a toxicant is lipid-soluble, it readily penetrates cell membranes. Many toxicants are stored in the body. Fat tissue, liver, kidney, and bone are the most common storage sites. Blood serves as the main avenue for distribution. Lymph also distributes some materials.

Excretion

The site and rate of **excretion** is another major factor affecting the toxicity of a xenobiotic. The kidney is the primary excretory organ, followed by the gastrointestinal tract, and the lungs (for gases). Xenobiotics may also be excreted in sweat, tears, and milk.

A large volume of blood serum is filtered through the kidney. Lipid-soluble toxicants are reabsorbed and concentrated in kidney cells. Impaired kidney function causes slower elimination of toxicants and increases their toxic potential.

Health Status

The health of an individual or organism can play a major role in determining the levels and types of potential toxicity. For example, an individual may have pre-existing kidney or liver disease. Certain conditions, such as pregnancy, also are associated with physiological changes in kidney function that could influence toxicity.

Nutritional Status

Diet (nutritional status) can be a major factor in determining who does or does not develop toxicity. For example:

- Consumption of fish that have absorbed mercury from contaminated water can result in mercury toxicity; an antagonist for mercury toxicity is the nutrient selenium.
- Some vegetables can accumulate cadmium from contaminated soil; an antagonist for cadmium toxicity is the nutrient zinc.
- Grapefruit contains a substance that inhibits the P450 drug detoxification pathway, making some drugs more toxic.

Find out more about nutrition and chemical toxicity **here**.

Circadian Rhythms

Circadian rhythms can play a role in toxicity. For example, rats administered an immunosuppressive drug had severe toxicity in their intestines 7 hours after light onset compared to controls and to other times in the day. The rats had changes in their digestive enzyme activity and other physiological indicators at this dosing time.

Find out more about circadian rhythm and gut toxicity **here**.

Other Factors

Presence of Other Chemicals

The presence of other chemicals, at the same time, earlier, or later may:

- **Decrease toxicity** (antagonism)
- **Add to toxicity** (additivity)
- **Increase toxicity** (synergism or potentiation)

For example:

- Antidotes used to counteract the effects of poisons function through antagonism (atropine counteracts poisoning by organophosphate insecticides).
- Alcohol may enhance the effect of many antihistamines and sedatives.
- A synergistic interaction between the antioxidant butylated hydroxytoluene (BHT) and a certain concentration of oxygen results in lung damage in the form of interstitial pulmonary fibrosis.

Information on additional examples of lung damage from chemical interactions can be found **here**.

Knowledge Check

1. A target organ is an organ that:

- ◯ Absorbs a toxic substance
- ◯ Stores an absorbed substance or its metabolite
- ◯ Is damaged by a toxic substance

Answer

Is damaged by a toxic substance

A target organ is an organ in which a substance exerts a toxic effect.

2. What are the important factors that influence the degree of toxicity of a substance?

- ◯ Innate chemical activity, form, dosage, and exposure route
- ◯ The species, life stage, gender, health status, nutritional status, and circadian rhythms of the organism
- ◯ Absorption, metabolism, distribution within the body, excretion, and presence of other chemicals
- ◯ All of the above

Answer

All of the above

3. Metabolism, or biotransformation, of a xenobiotic:

- ◯ Always results in reduced toxicity of the xenobiotic
- ◯ May result in detoxification or bioactivation
- ◯ Has no influence on the toxicity of the xenobiotic

Answer

May result in detoxification or bioactivation

Metabolism of a xenobiotic results in either detoxification, which converts the xenobiotic to a less toxic form, or bioactivation, which converts the xenobiotic to more reactive or toxic forms. For example, a xenobiotic itself might not be carcinogenic, but a metabolite of the xenobiotic might be.

4. An antibiotic administered to humans kills bacteria in the body but does not harm human tissues. This is an example of:

- ◯ Selective toxicity
- ◯ Acute toxicity
- ◯ Varying absorption of the antibiotic

Answer

Selective Toxicity

Selective toxicity refers to differences in toxicity between two species simultaneously exposed, much like the antibiotic in this example.

5. A major determinant of whether a toxicant will damage cells is its:

- ◯ Acidity
- ◯ Biotransformation
- ◯ Lipid solubility

Answer

Lipid solubility

A major determinant of whether or not a toxicant will damage cells is its lipid solubility. If a toxicant is lipid-soluble, it readily penetrates cell membranes.

This page titled 2.6.6.3.2: Factors Affecting Toxicity is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **3.2: Factors Affecting Toxicity** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.3.3: Systemic Toxic Effects

Types of Systemic Toxic Effects

Toxic effects are generally categorized according to the site of the toxic effect. In some cases, the effect may occur at only one site. This site is termed the **specific target organ**.

In other cases, toxic effects may occur at multiple sites. This is known as **systemic toxicity**. Types of systemic toxicity include:

- Acute Toxicity
- Subchronic Toxicity
- Chronic Toxicity
- Carcinogenicity
- Developmental Toxicity
- Genetic Toxicity (somatic cells)

Acute Toxicity

Acute toxicityoccurs almost immediately (seconds/minutes/hours/days) after an exposure. An **acute exposure** is usually a single dose or a series of doses received within a 24-hour period. Death can be a major concern in cases of acute exposures. For example:

- In 1989, 5,000 people died and 30,000 were permanently disabled due to exposure to methyl isocyanate from an industrial accident in India.
- Many people die each year from inhaling carbon monoxide from faulty heaters.


Close-up photo of the burner on a gas heater

Figure 2.6.6.3.3.1. Faulty gas heaters can emit toxic carbon monoxide
(Image Source: iStock Photos, ©)

Subchronic Toxicity

Subchronic toxicityresults from repeated exposure for several weeks or months. This is a common human exposure pattern for some pharmaceuticals and environmental agents. For example:

- Ingestion of warfarin (Coumadin®) tablets (blood thinners) for several weeks as a treatment for venous thrombosis can cause internal bleeding.
- Workplace exposure to lead over a period of several weeks can result in anemia.

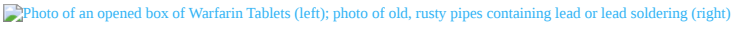

Photo of an opened box of Warfarin Tablets (left); photo of old, rusty pipes containing lead or lead soldering (right)

Figure 2.6.6.3.3.2. Warfarin Tablets (left); old lead pipes (right)
(Image Source: iStock Photos, ©)

Chronic Toxicity

Chronic toxicity represents cumulative damage to specific organ systems and takes many months or years to become a recognizable clinical disease. Damage due to subclinical individual exposures may go unnoticed. With repeated exposures or long-term continual exposure, the damage from this type of exposure slowly builds up (cumulative damage) until the damage exceeds the threshold for chronic toxicity. Ultimately, the damage becomes so severe that the organ can no longer function normally and a variety of chronic toxic effects may result.

Chronic toxic effects include:

- Cirrhosis in alcoholics who have ingested ethanol for several years.
- Chronic kidney disease in workmen with several years of exposure to lead.
- Chronic bronchitis in long-term cigarette smokers.
- Pulmonary fibrosis in coal miners (black lung disease).


Photo of cigarettes and liquor

Figure 2.6.6.3.3.3. Smoking cigarettes and/or drinking alcohol over a long period of time can lead to chronic toxicity
(Image Source: iStock Photos, ©)

Carcinogenicity

Carcinogenicity is a **complex multistage process** of abnormal cell growth and differentiation that can lead to cancer. The two stages of carcinogenicity are:

1. **Initiation** — a normal cell undergoes irreversible changes.
2. **Promotion** — initiated cells are stimulated to progress to cancer.

Chemicals can act as **initiators** or **promoters**.

The initial transformation that causes normal cells to undergo irreversible changes results from the mutation of the cellular genes that control normal cell functions. The mutation may lead to abnormal cell growth. It may involve a loss of suppresser genes that usually restrict abnormal cell growth. Many other factors are involved, such as growth factors, immune suppression, and hormones.

A **tumor (neoplasm)** is simply an uncontrolled growth of cells:

- **Benign tumors** grow at the site of origin; do not invade adjacent tissues or metastasize; and generally are treatable.
- **Malignant tumors (cancer)** invade adjacent tissues or migrate to distant sites **(metastasis)**. They are more difficult to treat and often cause death.

Developmental Toxicity

Developmental toxicity pertains to adverse toxic effects to the developing embryo or fetus. It can result from toxicant exposure to either parent before conception or to the mother and her developing embryo or fetus. The three basic types of developmental toxicity are:

1. **Embryolethality** — failure to conceive, spontaneous abortion, or stillbirth.
2. **Embryotoxicity** — growth retardation or delayed growth of specific organ systems.
3. **Teratogenicity** — irreversible conditions that leave permanent birth defects in live offspring, such as cleft palette or missing limbs.

Chemicals cause developmental toxicity in two ways:

1. They act directly on cells of the embryo, causing cell death or cell damage, leading to abnormal organ development.
2. They induce a mutation in a parent's germ cell, which is transmitted to the fertilized ovum. Some mutated fertilized ova develop into abnormal embryos.

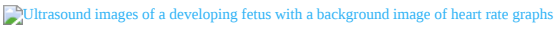

Ultrasound images of a developing fetus with a background image of heart rate graphs

Figure 2.6.6.3.3.4. Ultrasound images of a developing fetus
(Image Source: iStock Photos, ©)

Genetic Toxicity

Genetic toxicityresults from damage to DNA and altered genetic expression. This process is known as **mutagenesis**. The genetic change is referred to as a **mutation**and the agent causing the change is called a **mutagen**. There are three types of genetic changes:

1. **Gene mutation** — change in DNA sequence within a gene.
2. **Chromosome aberration** — changes in the chromosome structure.
3. **Aneuploidy** or **polyploidy** — increase or decrease in number of chromosomes.

If the mutation occurs in a germ cell, the effect is **heritable**. This means there is no effect on the exposed person; rather, the effect is passed on to future generations.

If the mutation occurs in a **somatic**cell, it can cause altered cell growth (for example, cancer) or cell death (for example, teratogenesis) in the exposed person.


Photo of a lab technician holding a vile, with a computer monitor displaying various genetic sequences in the background

Figure 2.6.6.3.3.5. Genetic toxicity results from damage to DNA and altered genetic expression
(Image Source: iStock Photos, ©)

This page titled 2.6.6.3.3: Systemic Toxic Effects is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **3.3: Systemic Toxic Effects** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.3.4: Organ Specific Toxic Effects

Organ Specific Toxic Effects

Toxic effects that pertain to specific organs and organ systems include:

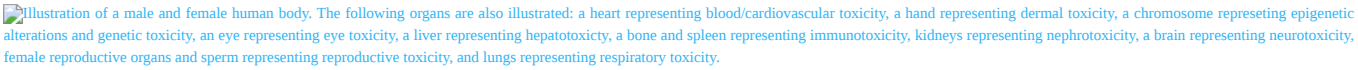
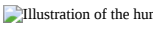

Illustration of a male and female human body. The following organs are also illustrated: a heart representing blood/cardiovascular toxicity, a hand representing dermal toxicity, a chromosome represeting epigenetic alterations and genetic toxicity, an eye representing eye toxicity, a liver representing hepatotoxicty, a bone and spleen representing immunotoxicity, kidneys representing nephrotoxicity, a brain representing neurotoxicity, female reproductive organs and sperm representing reproductive toxicity, and lungs representing respiratory toxicity.

Figure 2.6.6.3.4.1. Organ-specific toxic effects pertain to specific organs and organ systems
(Image Source: Adapted from iStock Photos, ©)

Blood and Cardiovascular/Cardiac Toxicity

Blood and Cardiovascular/Cardiac Toxicity results from xenobiotics acting directly on cells in circulating blood, bone marrow, and the heart. Examples of blood and cardiovascular/cardiac toxicity are:

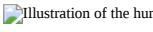
- Hypoxia due to carbon monoxide binding of hemoglobin preventing transport of oxygen.
- Decrease in circulating leukocytes due to chloramphenicol damage to bone marrow cells.
- Leukemia due to benzene damage of bone marrow cells.
- Arteriosclerosis due to cholesterol accumulation in arteries and veins.
- Death of normal cells in and around the heart as a result of exposure to **drugs used to treat cancer**.


Illustration of the human heartFigure 2.6.6.3.4.2. Heart cells can be damaged by exposure to certain drugs (Image Source: iStock Photos, ©)

Dermal Toxicity

Dermal Toxicity can occur when a toxicant comes into direct contact with the skin or is distributed to it internally. Effects range from mild irritation to severe changes, such as irreversible damage, hypersensitivity, and skin cancer. Examples of dermal toxicity include:

- Dermal irritation from skin exposure to gasoline.
- Dermal corrosion from skin exposure to sodium hydroxide (lye).
- Dermal itching, irritation, and sometimes painful **rash** from poison ivy, caused by **urushiol**.
- Skin cancer due to ingestion of arsenic or skin exposure to UV light.

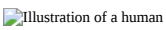

Illustration of the human handFigure 2.6.6.3.4.3. Hand (Image Source: iStock Photos, ©)

Epigenetic Alterations

Epigenetics is an emerging area in toxicology. In the field of genetics, epigenetics involves studying how external or environmental factors can switch genes on and off and change the programming of cells.

More specifically, epigenetics refers to stable changes in the programming of gene expression which can alter the phenotype without changing the DNA sequence (genotype). Epigenetic modifications include DNA methylation, covalent modifications of histone tails, and regulation by non-coding RNAs, among others.

Toxicants are examples of factors that can alter genetic programming.


Illustration of a human chromosomeFigure 2.6.6.3.4.4. Chromosome, which contains DNA (Image Source: iStock Photos, ©)

In the past, toxicology studies have assessed toxicity without measuring its impact at the level where gene expression occurs. Exogenous agents could cause long-term toxicity that continues after the initial exposure has disappeared, and such toxicities remain undetected by current screening methods. Thus, a current challenge in toxicology is to **develop screening methods that would detect epigenetic alterations** caused by toxicants.

Research is being done to assess epigenetic changes caused by toxicants. For example, the **National Institutes of Health (NIH) National Institute of Environmental Health Sciences (NIEHS) Environmental Epigenetics program** provides funding for a variety of research projects that use state-of-the-art technologies to analyze epigenetic changes caused by environmental exposures. NIEHS-supported researchers use animals, cell cultures, and human tissue samples to pinpoint how epigenetic changes can lead to harmful health effects and can potentially be passed down to the next generation.

Eye Toxicity

Eye Toxicity results from direct contact with or internal distribution to the eye. Because the cornea and conjunctiva are directly exposed to toxicants, conjunctivitis and corneal erosion may be observed following occupational exposure to chemicals. Many household items can cause conjunctivitis. Chemicals in the circulatory system can distribute to the eye and cause corneal opacity, cataracts, and retinal and optic nerve damage. For example:

- Acids and strong alkalis may cause severe corneal corrosion.
- Corticosteroids may cause cataracts.
- Methanol (wood alcohol) may damage the optic nerve.

Figure 2.6.6.3.4.5. *Eye (Image Source: iStock Photos, ©)*

Hepatotoxicity

Hepatotoxicity is toxicity to the liver, bile duct, and gall bladder. Because of its extensive blood supply and significant role in metabolism, the liver is particularly susceptible to xenobiotics Thus, it is exposed to high doses of the toxicant or its toxic metabolites. The primary forms of hepatotoxicity are:

- **Steatosis** — lipid accumulation in the hepatocytes.
- **Chemical hepatitis** — inflammation of the liver.
- **Hepatic necrosis** — death of the hepatocytes.
- **Intrahepatic cholestasis** — backup of bile salts into the liver cells.
- **Hepatic cancer** — cancer of the liver.
- **Cirrhosis** — chronic fibrosis, often due to alcohol.
- **Hypersensitivity** — immune reaction resulting in hepatic necrosis.

Figure 2.6.6.3.4.6. *Liver (Image Source: iStock Photos, ©)*

Related Resource: LiverTox®

Immunotoxicity

Immunotoxicity is toxicity of the immune system. It can take several forms:

- Hypersensitivity (allergy and autoimmunity)
- Immunodeficiency
- Uncontrolled proliferation (leukemia and lymphoma)

The normal function of the immune system is to recognize and defend against foreign invaders. This is accomplished by production of cells that engulf and destroy the invaders or by antibodies that inactivate foreign material. Examples include:

- Contact dermatitis due to exposure to poison ivy.
- Systemic lupus erythematosus ("lupus") in workers exposed to hydrazine.
- Immunosuppression by cocaine.
- Leukemia induced by benzene.



Figure 2.6.6.3.4.7. *Bone (which contains bone marrow) and spleen, both components of the immune system, which recognizes and defends against foreign invaders*
(Image Source: iStock Photos, ©)

Nephrotoxicity

The kidney is highly susceptible to toxicants because a high volume of blood flows through the organ and it filters large amounts of toxins which can concentrate in the kidney tubules.

Nephrotoxicity is toxicity to the kidneys. It can result in systemic toxicity causing:

- Decreased ability to excrete body wastes.

- Inability to maintain body fluid and electrolyte balance.
- Decreased synthesis of essential hormones (for example, erythropoietin, which increases the rate of blood cell production).

Figure 2.6.6.3.4.8. *Kidneys (Image Source: iStock Photos, ©)*

Neurotoxicity

Neurotoxicityrepresents toxicant damage to cells of the central nervous system (brain and spinal cord) and the peripheral nervous system (nerves outside the CNS). The primary types of neurotoxicity are:

- Neuronopathies (neuron injury)
- Axonopathies (axon injury)
- Demyelination (loss of axon insulation)
- Interference with neurotransmission

Figure 2.6.6.3.4.9. *Brain and synapse are susceptible to toxicant damage (Image Source: iStock Photos, ©)*

Reproductive Toxicity

Reproductive Toxicity involves toxicant damage to either the male or female reproductive system. Toxic effects may cause:

- Decreased libido and impotence.
- Infertility.
- Interrupted pregnancy (abortion, fetal death, or premature delivery).
- Infant death or childhood morbidity.
- Altered sex ratio and multiple births.
- Chromosome abnormalities and birth defects.
- Childhood cancer.

Figure 2.6.6.3.4.10. *Female reproductive organs (left); male and female germ cells (right) (Image Source: iStock Photos, ©)*

Respiratory Toxicity

Respiratory Toxicity relates to effects on the upper respiratory system (nose, pharynx, larynx, and trachea) and the lower respiratory system (bronchi, bronchioles, and lung alveoli). The primary types of respiratory toxicity are:

- Pulmonary irritation
- Asthma/bronchitis
- Reactive airway disease
- Emphysema
- Allergic alveolitis
- Fibrotic lung disease
- Pneumoconiosis
- Lung cancer

Figure 2.6.6.3.4.11. *Lungs (Image Source: iStock Photos, ©)*

Knowledge Check

Knowledge Check

1. Toxic effects are primarily categorized into two general types:

- ◯ Systemic or organ-specific effects
- ◯ Carcinogenic or teratogenic effects
- ◯ Hepatic or nephrotoxic effects

Answer

Systemic or organ-specific effects - **This is the correct answer.**
Toxic effects are broadly categorized as either systemic or organ- specific effects.

2. What is the main difference between acute and chronic toxicity?

- ◯ Different organs are involved
- ◯ Acute toxicity occurs only after a single dose, whereas chronic toxicity occurs with multiple doses
- ◯ Acute toxicity appears within hours or days of an exposure, whereas chronic toxicity takes many months or years to become a recognizable clinical disease
- ◯ Acute toxicity is less likely to lead to death than is chronic toxicity

Answer

Acute toxicity appears within hours or days of an exposure, whereas chronic toxicity takes many months or years to become a recognizable clinical disease - **This is the correct answer.**

3. Police respond to a 911 call in which two people are found dead in an enclosed bedroom heated by an unvented kerosene stove. There was no sign of trauma or violence. A likely cause of death is:

- ◯ Excess oxygen generated by the combustion of kerosene
- ◯ Acute toxicity due to uncombusted kerosene fumes
- ◯ Acute toxicity due to carbon monoxide poisoning

Answer

Acute toxicity due to carbon monoxide poisoning - **This is the correct answer.**
The victims most likely died as a result of acute toxicity from exposure to carbon monoxide.

4. Genetic toxicity can result in:

- ◯ Gene mutation
- ◯ Changes in the structure and/or number of chromosomes
- ◯ Epigenetic alterations
- ◯ All of the above

Answer

All of the above - **This is the correct answer.**
Genetic toxicity can cause gene mutations, changes in chromosome structure (aberration), increases or decreases in the number of chromosomes (aneuploidy or polyploidy), and changes to genetic programming (epigenetic alterations).

This page titled 2.6.6.3.4: Organ Specific Toxic Effects is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **3.4: Organ Specific Toxic Effects** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.4: Interactions

Learning Objectives

After completing this lesson, you will be able to:

- Explain the impacts that can be experienced when two or more chemicals interact.
- List realistic examples of chemical interactions.

In this section...

Topics include:

2.6.6.4.1: Interactions

Did you know?

Gasoline is a volatile, complex mixture of hydrocarbon compounds. The mixture is easily vaporized during handling in normal conditions. People are exposed to this complex substance during refueling at service stations. **More information** is available on consumer exposure to gasoline.

In this section, we will look into the effects of interactions among such chemicals.



Figure 1. Refueling car
(Image Source: iStock Photos, ©)

What We've Covered

In this section, we explored the following key points:

- Interactions between multiple chemicals can:
  - Decrease toxicity (antagonism).
  - Add to toxicity (additivity).
  - Increase toxicity (synergism or potentiation).
- Interactions can occur by simultaneous exposure or if exposure to the agents is separated by time.
- People are normally exposed to many chemicals and combinations of chemicals every day.
- Emerging approaches in assessing interactions include:
  - Adverse outcome pathways (AOPs).
  - *In vitro* methods.
  - "Omics" techniques.
  - *In silico* approaches.

Coming Up...

In the next section, we will explore various methods for testing toxicity.

This page titled 2.6.6.4: Interactions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.6.4.1: Interactions

Interactions

As noted in "**Factors Influencing Toxicity**," the presence of other chemicals, at the same time, earlier, or later may:

- Decrease toxicity (**antagonism**).
- Add to toxicity (**additivity**).
- Increase toxicity (**synergism** or **potentiation**)of some chemicals.

For example, interactions between two or more toxic agents can produce lung damage by interactions:

- Between chemicals.
- Between chemicals and receptors.
- In which a first agent modifies the cell and tissue response to a second agent.

Interactions may occur by:

- Simultaneous exposure to two or more agents.
- Exposure to two or more agents at different times.



Figure 2.6.6.4.1.1. Interactions between two or more substances often impact toxicity
(Image Source: iStock Photos, ©)

Sources of Interactions

Humans are normally exposed to many chemicals at one time. For example, the use of consumer products, medical treatments, and exposures from the diet and environment (such as from soil, air, and water) can consist of exposures to hundreds, if not thousands, of chemicals. Other examples include:

- Hospital patients receive an average of six drugs daily.
- Consumers may use five or more consumer products before breakfast (for example, soap, shampoo, conditioner, toothpaste, and deodorant).
- Home influenza treatment consists of aspirin, antihistamines, and cough syrup taken simultaneously.
- Drinking water may contain small amounts of pesticides, heavy metals, solvents, and other organic chemicals.
- Air often contains mixtures of hundreds of chemicals such as automobile exhaust and cigarette smoke.
- Gasoline vapor at service stations is a mixture of 40-50 chemicals.



Figure 2.6.6.4.1.2. The use of personal care products can result in expsoures to hundreds of chemicals
(Image Source: iStock Photos, ©)



Figure 2.6.6.4.1.3. *Cold and flu remedies are another source of chemical exposure*
(Image Source: iStock Photos, ©)

Toxicology studies and human health risk assessments have traditionally focused primarily on a single chemical. However, as noted above, people are exposed to many chemicals every day. They are also exposed to non-chemical stressors every day and throughout a lifetime.

In addition, non-chemical stressors include infectious agents, diet, and psychosocial stress, all of which have potential roles in contributing to the health effects associated with chemical exposures.

Approaches for Assessing Interactions

Development of methods to assess the health effects associated with complex exposures is underway at **various organizations**.

Non-animal tools and approaches are demonstrating high potential for use in assessing combined effects of chemicals on humans and the environment. These tools and approaches may help uncover information about new mixture components or entire mixtures, which can promote understanding of the underlying mechanisms of their combined effects. The strategies for assessing interactions rely less on *in vivo* testing and more on non-animal studies and computational tools and incorporate emerging approaches such as:

- The adverse outcome pathway (AOP) concept.
- *In vitro* methods.
- "Omics" techniques.
- *In silico* approaches such as quantitative structure activity relationships (QSARs).
- Read-across.
- Toxicokinetic modeling.
- Integrated approaches to testing and assessment (IATA).

The goals include the ability to development more effective and comprehensive regulatory assessments while reducing the reliance on animal testing.


Photo of a lab scientist typing on a computer keyboard, with a microscope and chemical flasks in the background

Figure 2.6.6.4.1.4. *Modern testing methods rely heavily on computational toxicology (Image Source: iStock Photos, ©)*

Knowledge Check

The presence of one chemical decreasing toxicity of another chemical is called:

- ◯ Additivity
- ◯ Antagonism
- ◯ Synergism

Answer

Antagonism - **This is the correct answer.**

When the presence of another chemical decreases toxicity of a chemical, this is called antagonism.

Additivity is when the combined toxic effect of two chemicals when given together is less than the sum of their individually measured toxic effects.

- ◯ True
- ◯ False

Answer

False - **This is the correct answer.**
An additive effect occurs when the combined effects of two or more chemicals is equal to the sum of the effects of each chemical given alone.

Piperonyl butoxide is not an insecticide; however, it can greatly increase the effects of a pyrethrum insecticide. Thus, piperonyl butoxide can be called a synergist and this interaction can be called synergism.

- ◯ True
- ◯ False

Answer

True - **This is the correct answer.**
The interaction of this combination is synergism. Synergists are used to enhance the toxicity of several commonly used insecticides.

This page titled 2.6.6.4.1: Interactions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **4.1: Interactions** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.5: Toxicity Testing Methods

Learning Objectives

After completing this lesson, you will be able to:

- Explain modern approaches to testing for and assessing toxicity.
- Identify sources of information related to alternatives to using animals to assess toxicity.
- Explain how clinical investigations and epidemiology studies are used to evaluate toxicity to humans.

In this section...

Topics include:

2.6.6.5.1: Testing and Assessing Toxicity

2.6.6.5.2: Clinical Investigations and Other Types of Human Data

2.6.6.5.3: Epidemiology Studies

What We've Covered

This section included the following key points:

- The 3Rs concept of using test methods replace the use of animals with other types of studies and approaches, reduce the number of animals used in studies, and refine study procedures to cause less pain or stress to animals.
- ALTBIB is a comprehensive starting point provided by NLM to find information related to alternatives to animal testing.
- Animal tests for toxicity have been conducted prior to and in parallel with human clinical investigations.
- Standardized animal tests have been developed for testing:
  - Acute toxicity
  - Subchronic toxicity
  - Chronic toxicity
  - Carcinogenicity
  - Reproductive toxicity
  - Developmental toxicity
  - Dermal toxicity
  - Ocular toxicity
  - Neurotoxicity
  - Genetic toxicity
- Modern approaches to toxicity testing are preferred over animal testing and include:
  - *In vitro* methods, which are performed outside living organisms.
  - *In silico* methods, which are performed using computers and computer simulation.
  - Chip models, which include human cell cultures placed on computer chips for study.
- Approaches used for testing pharmaceuticals include:
  - Clinical investigations, in which human subjects are studied with clinical observations and laboratory measurements.
  - Epidemiological studies, involving observation of humans exposed to xenobiotics in their regular life or occupation.
  - Reports of adverse reactions to drugs.
- Consumer products and the chemicals they contain are tested through:
  - *In silico* data from computer models.

- *In vitro* data from tests performed as alternatives to animal testing.
- Animal study data.
- Human data from premarketing and postmarketing studies.

Coming Up...

In the next section, we will explore the concept of risk assessment.

This page titled 2.6.6.5: Toxicity Testing Methods is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.6.5.1: Testing and Assessing Toxicity

Testing and Assessing Toxicity

Alternatives to animal testing have emerged in recent years.

Since about 1990, numerous attempts have been made around the world to reduce the use of and replace laboratory animals in toxicology and other studies. These efforts have involved finding alternatives to animal testing and incorporating the "3Rs" concept (Replace, Reduce and Refine), which means using test methods that:

- **Replace** the use of animals with other types of studies and approaches.
- **Reduce** the number of animals in studies.
- **Refine** the procedures to make studies less painful or stressful to the animals.

Regulatory authorities, companies, and others have endorsed the principle of the 3Rs, and alternative testing methods have been and are being developed. An international group that has played a key role is the International Cooperation on Alternative Test Methods (ICATM). Established in 2009, ICATM includes representatives of **organizations from various countries**.

Finding Information about Alternatives to Animal Testing

Many countries including the United States, Canada, and the European Union member states, require that a comprehensive search for possible alternatives be completed before some or all research involving animals is begun. Because numerous Web resources are now available to provide guidance and other information on *in vitro* and other alternatives to animal testing, completing such searches and keeping current with information associated with alternatives to animal testing is much easier than it used to be.

ALTBIB, from NLM

The NLM ALTBIB ("Resources for Alternatives to the Use of Live Vertebrates in Biomedical Research and Testing") portal is a comprehensive starting point for finding information related to alternatives to animal testing. ALTBIB is available at **http://toxnet.nlm.nih.gov/altbib.html**.

It provides access to **PubMed®/MEDLINE®** citations relevant to alternatives to use of live vertebrates in biomedical research and testing.

ALTBIB's topics and subtopics are aligned with current approaches. For example, information is provided on *in silico*, *in vitro*, and improved (refined) animal testing methods and on testing strategies that incorporate these methods and other approaches.

Screenshot of the ALTBIB homepageFigure 2.6.6.5.1.1. NLM ALTBIB homepage

ALTBIB also provides access to news and additional resources, including information on the status of the evaluation and acceptance of alternative methods. Main categories include:

- Animal Alternatives News
- Additional Resources
- Evaluation/Acceptance of Test Methods

Links to Specific Resources (Sources Providing Animal Alternatives News, Key Organizations Providing Resources, and the Regulatory Acceptance of Specific Alternative Methods and Milestones in Non-animal Toxicity Testing)

Animal Tests

NOTE: This information is provided for historical and other reasons, especially since animal testing is still being done in some cases, and because toxicologists, risk assessors, and others are faced with interpreting the results of new and old studies that used animals.

Animal tests for toxicity have been conducted prior to and in parallel with human clinical investigations as part of the non-clinical laboratory tests of pharmaceuticals. For pesticides and industrial chemicals, human testing is rarely conducted. Years ago, results from animal tests were often the only way to effectively predict toxicity in humans.

Animal tests were developed and used because:

- Chemical exposure can be precisely controlled.
- Environmental conditions can be well-controlled.
- Virtually any type of toxic effect can be evaluated.
- The mechanism by which toxicity occurs can be studied.



Figure 2.6.6.5.1.2. *Rats have traditionally been used in toxicity studies using animals (Image Source: iStock Photos, ©)*

Standardized Animal Toxicity Tests

Animal methods to evaluate toxicity have been developed for a wide variety of toxic effects. Some procedures for routine safety testing have been standardized. **Standardized animal toxicity**tests have been highly effective in detecting toxicity that may occur in humans. As noted above, concern for animal welfare has resulted in tests that use humane procedures and only as many animals as are needed for statistical reliability.

To be standardized, a test procedure must have scientific acceptance as the most meaningful assay for the toxic effect. Toxicity testing can be very **specific** for a particular effect, such as dermal irritation, or it may be **general**, such as testing for unknown chronic effects.

Standardized tests have been developed for the following effects:

- Acute Toxicity
- Subchronic Toxicity
- Chronic Toxicity
- Carcinogenicity
- Reproductive Toxicity
- Developmental Toxicity
- Dermal Toxicity
- Ocular Toxicity
- Neurotoxicity
- Genetic Toxicity

Species Selection

Species selectionvaries with the toxicity test to be performed. There is no single species of animal that can be used for all toxicity tests. Different species may be needed to assess different types of toxicity. The published literature (such as via PubMed) and online databases (such as TOXNET) should be searched for information from non-animal and animal studies, as well as for possible best approaches, most applicable species, and strains and gender of a species. Here are two examples:

- It would have been invaluable years ago for toxicologists and risk assessors to have known that carcinogenic effects in male rats are considered irrelevant for humans if the α(2u)-globulin protein is involved because humans lack that protein. See **another example**
- Many physiological, pharmacological, and toxicological findings related to organic anion and cation transport and transporters in rodents and rabbits do not apply to humans. **Learn more**

In some cases, it may not be possible to use the most desirable animal for testing because of animal welfare or cost considerations.

- For example, use of dogs and non-human primates is now restricted to special cases or banned by some organizations, even though they represent the species that may respond the closest to humans in terms of chemical and other exposures (however, note the examples above).

Rodents and rabbits are the most commonly used laboratory species because they are readily available, inexpensive to breed and house, and they have a history of producing reliable results in experiments.

The toxicologist attempts to design an experiment to duplicate the potential exposure of humans as closely as possible. For example:

- The **route of exposure**should simulate that of human exposure. Most standard tests use inhalation, oral, or dermal routes of exposure.
- The **age of test animals**should relate to that of humans. Testing is normally conducted with young adults, although in some cases, newborn or pregnant animals may be used.
- For most routine tests, **both sexes**are used. Sex differences in toxic response are usually minimal, except for toxic substances with hormonal properties.
- **Dose levels**are normally selected so as to determine the threshold as well as a dose-response relationship. Usually, a minimum of three dose levels are used.



Figure 2.6.6.5.1.3. *Rodents have commonly been used in animal testing (Image Source: iStock Photos, ©)*

Acute Toxicity

Historically, **acute toxicity tests** were the first tests conducted. They provide data on the relative toxicity likely to arise from a single or brief exposure, or sometimes multiple doses over a brief period of time. Standardized tests are available for oral, dermal, and inhalation exposures, and many regulatory agencies still require the use of all or some of these tests. Table 2.6.6.5.1.1 lists basic parameters historically used in acute toxicity testing.

Category	Parameter
Species	Rats preferred for oral and inhalation tests; rabbits preferred for dermal tests
Age	Young adults
Number of animals	5 of each sex per dose level
Dosage	Three dose levels recommended; exposures are single doses or fractionated doses up to 24 hours for oral and dermal studies and 4-hour exposure for inhalation studies
Observation period	14 days

Table 2.6.6.5.1.1. Acute toxicity test parameter

Subchronic Toxicity

Subchronic toxicity testsare employed to determine toxicity likely to arise from repeated exposures of several weeks to several months. Standardized tests are available for oral, dermal, and inhalation exposures. Detailed information is obtained during and after the study, ranging from body weight, food and water consumption measurements, effects on eyes and behavior, composition of blood, and microscopic examination of selected tissues and organs.

Table 2.6.6.5.1.2 lists basic parameters previously used in subchronic toxicity testing.

Category	Parameter
Species	Rodents (usually rats) preferred for oral and inhalation studies; rabbits for dermal studies; non-rodents (usually dogs) recommended as a second species for oral tests
Age	Young adults
Number of animals	10 of each sex for rodents; 4 of each sex for non-rodents per dose level
Dosage	Three dose levels plus a control group; includes a toxic dose level plus NOAEL; exposures are 90 days
Observation period	90 days (same as treatment period)

Table 2.6.6.5.1.2. *Subchronic toxicity test parameter*

Chronic Toxicity

Chronic toxicity testsdetermine toxicity from exposure for a substantial portion of a subject's life. They are similar to the subchronic tests except that they extend over a longer period of time and involve larger groups of animals.

Table 2.6.6.5.1.3 includes basic parameters previously used in chronic toxicity testing.

Category	Parameter
Species	Two species recommended; rodent and non-rodent (rat and dog)
Age	Young adults
Number of animals	20 of each sex for rodents, 4 of each sex for non-rodents per dose level
Dosage	Three dose levels recommended; includes a toxic dose level plus NOAEL. The recommended maximum chronic testing durations for pharmaceuticals are now 6 and 9 months in rodents and non-rodents, respectively. (Historically exposures were for 12 months, 24 months for food chemicals.)
Observation period	12-24 months

Table 2.6.6.5.1.3. *Chronic toxicity test parameter*

Carcinogenicity

Carcinogenicity testsare similar to chronic toxicity tests. However, they extend over a longer period of time and require larger groups of animals in order to assess the potential for cancer.

Table 2.6.6.5.1.4 lists basic parameters used in the past in carcinogenicity testing.

Category	Parameter
Species	Testing in two rodents species—rats and mice—is preferable given their relatively short life spans.
Age	Young adults
Number of animals	50 of each sex per dose level
Dosage	Three dose levels recommended; highest should produce minimal toxicity; exposure periods are at least 18 months for mice and 24 months for rats
Observation period	18-24 months for mice and 24-30 months for rats

Table 4. Carcinogenicity test parameter

Reproductive Toxicity

Reproductive toxicity testing is intended to determine the effects of substances on gonadal function, conception, birth, and the growth and development of offspring. The oral route of administration is preferable.

Table 2.6.6.5.1.5 lists basic parameters historically used in reproductive toxicity testing.

Category	Parameter
Species	Rat is recommended
Age	Young adults
Number of animals	20 of each sex per dose level
Dosage	Three dose levels recommended; highest dose should produce toxicity but not mortality in parents; lowest dose should not produce toxicity
Observation period	Test substance given to parental animals (P1) prior to mating, during pregnancy, and through weaning of first generation (F1) offspring; substance then given to selected F1 offspring during their growth into adulthood, mating, and production of second generation (F2) until the F2 generation is 21 days old.

Table 2.6.6.5.1.5. *Reproductive toxicity test parameter*

Developmental Toxicity

Developmental toxicitytesting detects the potential for substances to produce embryotoxicity and birth defects.

Table 2.6.6.5.1.6 lists basic parameters previously used in developmental toxicity tests.

Category	Parameter
Species	Two species recommended; rat, mouse, hamster, and rabbit are used most commonly.
Age	Young adult females
Number of animals	20 pregnant females per dose level
Dosage	At least three dose levels are used; includes a toxic dose level plus NOAEL; occurs throughout organ development in the fetus for teratogenic effects; starts with parents prior to breeding, continues during pregnancy for all developmental effects
Observation period	Offspring sacrificed and examined the day prior to expected birth for teratogenic effects; offspring observed for growth retardation and abnormal function through infancy and examined for teratogenic effects

Table 2.6.6.5.1.6. *Developmental toxicity test parameter*

Dermal Toxicity

Dermal toxicity testsdetermine the potential for an agent to cause irritation and inflammation of the skin. Those reactions may be a result of direct damage to the skin cells by a substance or an indirect response due to sensitization from prior exposure. *In vitro* approaches to dermal toxicity testing are being developed, in part because this type of testing has received so much publicity.

Table 2.6.6.5.1.7 lists basic parameters historically used in dermal toxicity testing.

Category	Parameter
Primary dermal irritation	Determines direct toxicity. The substance is applied to the skin of 6 albino rabbits for 4 hours and the rabbits are observed for 72 hours for irritation.
Dermal sensitization	Assays for immune hypersensitivity of the skin, consisting of two phases: 1) Application of the test substance to the skin of guinea pigs for 4 hours in the sensitization phase; and 2) a challenge phase at least 1 week later in which the substance is reapplied to the skin. An inflammatory reaction indicates that the skin has been sensitized to the substance.

Table 2.6.6.5.1.7. *Dermal toxicity test parameter*

Ocular Toxicity

Ocular toxicity was at one time determined by applying a test substance for 1 second to the eyes of 6 test animals, usually rabbits. The eyes were then carefully examined for 72 hours, using a magnifying instrument to detect minor effects. An ocular reaction can occur on the cornea, conjunctiva, or iris. It may be simple irritation that is reversible and quickly disappears.

This eye irritation test was commonly known as the **"Draize Test."**This test has received much attention, such as the development of a **"low volume" variation and *in vitro* approaches**.

Neurotoxicity

A battery of **standardized neurotoxicity tests**were developed to supplement the **delayed neurotoxicity test**in domestic chickens (hens). The hen assay determines delayed neurotoxicity resulting from exposure to anticholinergic substances, such as certain pesticides. The hens are protected from immediate neurological effects of the test substance and observed for 21 days for delayed neurotoxicity.

Table 2.6.6.5.1.8 lists measurements included in other neurotoxicity tests.

Category	Parameter
Motor activity	Tests for decreased motor activity, such as cage movement. Rats or mice are used.
Peripheral nerve conduction	Tests for electrical conduction in motor and sensory nerves. Rodents are exposed to the test substance for 90 days.
Neuropathy	Tests for nerve damage by microscopic examination. This is one aspect of other standardized toxicity tests.

Table 2.6.6.5.1.8. Neurotoxicity test parameter

Genetic Toxicity

Genetic toxicityis determined using a wide range of test species including whole animals and plants *(for example, rodents, insects, and corn)*, microorganisms, and mammalian cells. A large variety of tests have been developed to measure gene mutations, chromosome changes, and DNA activity.

Table 2.6.6.5.1.9 lists parameters used for common gene mutation tests.

Category	Parameter
Microorganisms	*Salmonella typhimurium* and *Escherichia coli* are commonly used bacterial tests. The *S. typhimurium* assay is known as the Ames Test. Yeasts are also used to detect gene mutations in culture systems.
Mammalian cells	The two main cell lines are mouse lymphoma and Chinese hamster ovary (CHO) cells.
Fruit Flies	*Drosophila melanogaster* is used to detect sex-linked recessive lethal mutations.
Mice	The Mouse Specific-Locus Test is the major gene mutation test that employs whole animals. Exposed mice are bred and observed for hereditary changes.

Table 2.6.6.5.1.9. Genetic toxicity test parameter

Chromosomal effectscan be detected with a variety of tests, some of which utilize entire animals *(in vivo) and* some which use cell systems *(in vitro)*. Several assays are available to test for chemically induced chromosome aberrations in whole animals. Table 2.6.6.5.1.10 lists common *in vivo means of* testing chromosomal effects.

Category	Parameter
Rodent chromosomal assay	Involves exposure of mice or rats to a single dose of a substance. Their bone marrow is analyzed for chromosome aberrations over a 48-hour period.
Dominant lethal assay	Exposed male mice or rats are mated with untreated females. The presence of dead implants or fetuses is the result of the fertilized ovum receiving damaged DNA from the sperm. This leads to the death of the embryo or fetus. The genetic defect in the sperm is thus a heritable dominant lethal mutation.
Micronucleus test	Mice are exposed once and their bone marrow or peripheral blood cells are examined for 72 hours for the presence of micronuclei, such as broken pieces of chromosomes surrounded by a nuclear membrane.
Heritable translocation assay	Exposed male Drosophila or mice are bred to non-exposed females. The offspring males (F1 generation) are then bred to detect chromosomal translocations.
Sister chromatid exchange assay (SCE)	Mice are exposed to a substance and their bone marrow cells or lymphocytes are examined microscopically for complete chromosomal damage. This is indicated by chromatid fragments joining sister chromatids rather than their own.

Table 2.6.6.5.1.10. *Chromosomal effects (In vivo) test parameter*

In Vitro Testing

In vitro tests for chromosomal effects involve exposure of cell cultures and followed by microscopic examination of them for chromosome damage.

The most commonly used cell lines are Chinese Hamster Ovary (CHO) cells and human lymphocyte cells. The CHO cells are easy to culture, grow rapidly, and have a low chromosome number (22), which makes for easier identification of chromosome damage.

Human lymphocytes are more difficult to culture. They are obtained from healthy human donors with known medical histories. The results of these assays are potentially more relevant to determine effects of xenobiotics that induce mutations in humans.

Two widely used genotoxicity tests measure DNA damage and repair that is not mutagenicity. DNA damage is considered the first step in the process of mutagenesis. Common assays for detecting DNA damage include:

1. **Unscheduled DNA synthesis (UDS)** — involves exposure of mammalian cells in culture to a test substance. UDS is measured by the uptake of tritium-labeled thymidine into the DNA of the cells. Rat hepatocytes or human fibroblasts are the mammalian cell lines most commonly used.
2. **Exposure of repair-deficient *E. coli* or *B. subtilis*** — DNA damage cannot be repaired so the cells die or their growth may be inhibited.

Emerging Approaches and Methods

In the future, there will likely be additional and refined *in vitro* methods, and the emergence of *in silico*and "chip" approaches.

Many current efforts are underway to refine, develop, and validate *in vitro* methods.

Did you know?

The Human Toxicology Project Consortium provides a video series called **Pathways to a Better Future**. These videos discuss the future of toxicology, if you would like to know more about where the field is headed.

In Silico Methods

Also emerging are *in silico* methods, meaning "performed on computer or via computer simulation." This term was developed as an analogy to the Latin phrases *in vivo* and *in vitro.*

Advanced computer models called "Virtual Tissue Models" are being developed by the U.S. **EPA's National Center for Computational Toxicology (NCCT)**. The EPA's Virtual Tissue Models are described as using "new computational methods to construct advanced computer models capable of simulating how chemicals may affect human development. Virtual tissue models are some of the most advanced methods being developed today. The models will help reduce dependence on animal study data and provide much faster chemical risk assessments" (**source**).

One example is the Virtual Embryo (**v-Embryo™**) research effort, aimed at developing prediction models to increase our understanding of how chemical exposure may affect unborn children. Researchers are integrating new types of *in vitro*, *in vivo*, and *in silico* models that simulate critical steps in fetal development. Virtual Embryo models simulate biological interactions observed during development and predict chemical disruption of key biological events in pathways that is believed to lead to adverse effects.

Figure 2.6.6.5.1.4. *v-Embryo logo*

"Chip" Models

Also emerging are microphysiological systems (MPS) that are used in "tissue chip" and "organs-on-chips" models. Chip models include human cell cultures that are placed on a computer chip and studied there. The **Wyss (pronounced "Veese") Institute for Biologically Inspired Engineering** is a helpful resource for more information.

For example, the "Lung-on-a-chip" is described as "combining microfabrication techniques with modern tissue engineering, lung-on-a-chip offers a new *in vitro* approach to drug screening by mimicking the complicated mechanical and biochemical behaviors of a human lung." To learn more, watch a **video** from the Wyss Institute that shows a human lung-on-a-chip. **Another Wyss Institute video** illustrates how researches have used long-on-a-chip to mimic pulmonary edema.

Figure 2.6.6.5.1.5. *Lung-on-a-chip used to mimic pulmonary edema*
(Image Source: The Wyss Institute for Biologically Inspired Engineering)

Using a connected series of tissue chips as an integrated multi-organ system can allow for the creation of a "**human-on-a-chip**," to be used to model the metabolism and effects of drugs and other substances moving through a human. For example, a liver chip could provide fluids and metabolites to a kidney chip, allowing for the assessment of the nephrotoxic (kidney damage) potential of a substance metabolized in the liver.

Induced Pluripotent Stem Cells (iPSCs)

Induced pluripotent stem cells (iPSCs) are an emerging approach using *in vitro* cultures of cells. The cells of mammals and plants can be reprogrammed via "cellular reprogramming" to generate iPSCs. Like human embryonic stem cells, iPSCs are pluripotent (capable of giving rise to several different cell types) and these cells can renew themselves. As examples, iPSC-derived hepatocytes, cardiomyocytes, and neural cells can serve as tools for the screening of drugs and other substances for potential toxicity, and also can be used to study disease mechanisms and pathways. Further, iPSCs have been studied in immunotherapy and regenerative cellular therapies.

Figure 2.6.6.5.1.6. *Promise of hiPSCs. Schematic representation of how somatic cells taken from a patient can be reprogrammed into induced pluripotent stem cells (iPSCs) using the 'Yamanaka' factors, OCT4, KLF4, c-MYC and SOX2. Subsequent differentiation of human iPSCs (hiPSCs) into neurons of define lineage allow for investigations into disease pathophysiology and identification of potential drug targets. In addition, hiPSC derived neurons may function as a cellular platform in which drug screens can be carried out using disease relevant neurons.*
*(Image Source: **Adapted under Creative Commons Attribution License (CC BY)**.*
*doi: **10.1016/j.yhbeh.2015.06.014***
*Original image: **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4579404/figure/f0005/**)*

Did you know?

Professor Shinya Yamanaka (Kyoto University, Japan) received the 2012 Nobel Prize in Physiology or Medicine for discovering that mature cells can be reprogrammed to iPSCs that can differentiate into any type of cell. Key to this discovery was his use of four "reprogramming factors" referred to as c-Myc, Klf4, Oct3/4, and Sox2.

Learn more

Combining "Chips" and iPSCs

The emerging approaches of "chips" and iPSCs are being combined. One example is for the evaluation of drugs as potential countermeasures for biological and chemical threats that can be a substitute for human clinical trials. The "chips" and "humans on a chip" can be used as complex *in vitro* human models to simulate the biology and function of an organ.

This page titled 2.6.6.5.1: Testing and Assessing Toxicity is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **5.1: Testing and Assessing Toxicity** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.5.2: Clinical Investigations and Other Types of Human Data

For Drugs

The focus of this section is on the U.S. Food and Drug Administration (FDA), but regulatory agencies worldwide have very similar approaches. The main methods of determining the toxicity of drugs to humans are:

- **Clinical investigations** — administration of chemicals to human subjects with careful clinical observations and laboratory measurements.
- **Epidemiological studies** — observation of humans who have been exposed to xenobiotics in the normal course of their life or occupation.
- **Adverse reactions to drug reports** — reports voluntarily submitted by physicians to the FDA after a drug has been approved and is in widespread use.



Figure 2.6.6.5.2.1. Drugs can be toxic to humans
(Image Source: iStock Photos, ©)

Clinical Investigations

Clinical investigationsare a component of **Investigational New Drug Applications (INDs)** submitted to the FDA. Clinical investigations are conducted only after a minimal battery of nonclinical laboratory studies has been completed.

Toxicity studies using human subjects require strict ethical considerations. They are primarily conducted for new pharmaceutical applications submitted to the FDA for approval.

Generally, toxicity found in animal studies occurs with similar incidence and severity in humans. Differences sometimes occur, thus clinical tests with humans are needed to confirm the results of nonclinical laboratory studies.

FDA clinical investigations are conducted in three phases, as outlined below.



Figure 2.6.6.5.2.2. *Portion of the Investigational New Drug Application (IND)*
(Image Source: FDA)

Phase 1 consists of testing the drug in a small group of 20 to 80 healthy volunteers. Information obtained in Phase 1 studies is used to design Phase 2 studies, in particular, to determine the drug's:

- Initial tolerability in human subjects.
- Pharmacokinetics and pharmacological effects.



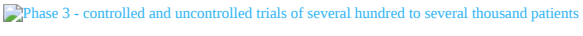
Phase 2 studies are more extensive, involving several hundred patients and are used to:

- Determine the short-term side effects of the drug.
- Determine the risks associated with the drug.
- Evaluate the effectiveness of the drug for treatment of a particular disease or condition.
- Elucidate the drug's metabolism.



Phase 3 studies are **controlled** and uncontrolled trials conducted with several hundred to several thousand patients. They are designed to:

- Gather additional information about effectiveness and safety.
- Evaluate overall risk:benefit profile of the drug.
- Provide the basis for the precautionary information that accompanies the drug.



For Consumer Products

Health-related data for a chemical in a consumer product (and for the consumer product itself for the human studies) can come from the following types of studies:

- *In silico* data — from computer programs that estimate toxic properties based on data for similar chemicals, and/or from the physical chemical properties.
- *In vitro* data — from the results of alternatives to animal tests, such as from cell cultures used to assess the potential for eye or skin irritation.
- Animal (toxicological) study data — for example, from studies that assessed eye or skin irritation potential.
- Human data – from studies conducted before (**premarketing**) and after (**postmarketing**) a product had been sold to consumers. More specifically, from:
  - Premarketing **clinical studies**, such as from patch tests to assess skin irritation potential.
  - Premarketing **"controlled use" studies** that are designed to assess the skin effects from using a new type of personal care product.
  - Postmarketing **studies conducted by physicians or dermatologists**, such as testing a diagnostic patch with their patients.
  - Postmarketing **epidemiological studies**, including studies developed by Poison Control Centers, companies, and academia that look at the "real world" health reports of effects associated with consumer use of a product.



Figure 2.6.6.5.2.3. *Toxicologists in a lab, using a computer for research*
(Image Source: iStock Photos, ©)

Did you know?

Bisphenol A (BPA) and **phthalates** are chemicals that have been widely found in consumer products. BPA has been used in some food can linings, polycarbonate food and beverage containers, tooth sealants applied to dentists, and even in cash register receipts! Examples of potential exposures to BPA include eating or drinking foods or liquids from those containers, and skin exposures from handing the cash register receipts. Workers involved in making products with BPA can be exposed during production.

Often called plasticizers, phthalates are used to make plastics more flexible. Some phthalates are used as solvents. They can be found in vinyl flooring and shower curtains, children's toys, personal care products, and as contaminants in the food supply. As with BPA, exposures can come from many sources.

Toxicologists and others are still assessing the full extent of the potential impacts on health. Studies suggest that BPA and phthalates affect the reproductive system, impacting how hormones such as estrogen and testosterone work in the body. The impact of fetal or early childhood exposures is still being assessed. Because of the ubiquity of the possible products containing these chemicals, thorough assessments of potential exposures, toxicities, and potential substitutes are essential.



Figure 2.6.6.5.2.4. *Plastic food and beverage containers are common sources of BPA and phthalates in consumer products*
(Image Source: iStock Photos, ©)

This page titled 2.6.6.5.2: Clinical Investigations and Other Types of Human Data is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **5.2: Clinical Investigations and Other Types of Human Data** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.5.3: Epidemiology Studies

What Are Epidemiology Studies?

Epidemiology studies are conducted using human populations to evaluate whether there is a **correlation or causal relationship** between exposure to a substance and adverse health effects.

These studies differ from clinical investigations in that individuals have already been administered the drug during medical treatment or have been exposed to it in the workplace or environment.

Epidemiological studies measure the risk of illness or death in an exposed population compared to that risk in an identical, unexposed population (for example, a population the same age, sex, race and social status as the exposed population).



Figure 2.6.6.5.3.1. Epidemiology studies tend to produce graphs and charts for data analysis and presentation
(Image Source: Adapted from iStock Photos, ©)

Types of Studies

There are four primary types of epidemiology studies. They are:

1. **Cohort studies** — A cohort (group) of individuals with exposure to a chemical and a cohort without exposure are followed over time to compare disease occurrence.
2. **Case control studies** — Individuals with a disease (such as cancer) are compared with similar individuals without the disease to determine if there is an association of the disease with prior exposure to an agent.
3. **Cross-sectional studies** — The prevalence of a disease or clinical parameter among one or more exposed groups is studied, such as:
    - The prevalence of respiratory conditions among furniture makers.
4. **Ecological studies** – The incidence of a disease in one geographical area is compared to that of another area, such as:
    - Cancer mortality in areas with hazardous waste sites as compared to similar areas without waste sites.

Cohort Studies

Cohort studies are the most commonly conducted epidemiology studies and they frequently involve occupational exposures. Exposed persons are easy to identify and their exposure levels are usually higher than in the general public. There are two types of cohort studies:

1. **Prospective**, in which cohorts are identified based on current exposures and followed into the future.
2. **Retrospective,** in which cohorts are identified based on past exposure conditions and study "follow-up" proceeds forward in time; data come from past records.

Common Statistical Measures

Standard, quantitative measures are used to determine if epidemiological data are meaningful. The most commonly used measures are:

- **Odds Ratio (O/R)** — The ratio of risk of disease in a case-control study for an exposed group to an unexposed group. An odds ratio equal to 2 (O/R = 2) means that the exposed group has twice the risk as the non-exposed group.
- **Standard Mortality Ratio (SMR)** — The relative risk of death based on a comparison of an exposed group to non-exposed group. A standard mortality ratio equal to 150 (SMR = 150) indicates that there is a 50% greater risk.
- **Relative Risk (RR)** — The ratio expressing the occurrence of disease in an exposed population to that of an unexposed population. A relative risk of 175 (RR = 175) indicates a 75% increase in risk.

Study Design

When designing an epidemiology study, the most critical aspects include:

- An appropriate control group.

- An adequate time span.
- The statistical ability to detect an effect.

More specifically, the control population used as a comparison group must be as similar as possible to that of the test group, for example, same age, sex, race, social status, geographical area, and environmental and lifestyle influences.

Many epidemiology studies evaluate the potential for an agent to cause cancer. Because most cancers require long latency periods, the study must cover that period of time.

The statistical ability to detect an effect is referred to as the **power** of the study. To gain precision, the study and control populations should be as large as possible.

Bias Errors

Epidemiologists attempt to control errors that can occur in the collection of data, which are known as bias errors. The three main types of bias errors are:

1. **Selection bias**, which occurs when the study group is not representative of the population from which it came.
2. **Information bias**, which occurs when study subjects are misclassified as to disease or exposure status. Recall bias occurs when individuals are asked to remember exposures or conditions that existed years before.
3. **Confounding factors**, which occur when the study and control populations differ with respect to factors which might influence the occurrence of the disease. For example, smoking might be a confounding factor and should be considered when designing studies.

Postmarketing Studies

Finally, for consumer products, **postmarketing** epidemiological studies can be performed. Examples include studies developed by Poison Control Centers, companies, academia, and other sources to look at the "real world" health reports of effects associated with consumer use of a product or article under reasonably foreseeable conditions.

Knowledge Check

While animal testing was historically the primary method used in testing for toxicity, modern testing methods prefer:

- ◯ *In silico*
- ◯ *In vitro*
- ◯ Refined animal testing
- ◯ All of the above

Answer

All of the above - **This is the correct answer.**
Modern approaches to testing for toxicity include *in silico*, *in vitro*, and improved (refined) animal testing.

In testing a pharmaceutical to comply with FDA requirements, the initial testing consists of:

- ◯ Clinical investigations
- ◯ Non-clinical laboratory studies
- ◯ Epidemiology studies
- ◯ All of the above

Answer

Non-clinical laboratory studies - **This is the correct answer.**
Investigational New Drug Applications (IND) require clinical investigations. Before clinical investigations begin, a minimal battery of non-clinical laboratory studies must be completed.

The primary goal of a Phase 1 clinical investigation is to:

- ◯ Obtain information to design a more definitive Phase 2 clinical investigation
- ◯ Evaluate the effectiveness of a drug for treating disease
- ◯ Provide the basis for pharmaceutical labeling

Answer

Obtain information to design a more definitive Phase 2 clinical investigation - **This is the correct answer.**
The primary goal of a Phase 1 clinical investigation is to obtain information that is used to design more extensive, Phase 2 studies.

Determining the overall risk versus the benefit of a new pharmaceutical is part of:

- ◯ Phase 2 clinical investigation
- ◯ Phase 3 clinical investigation
- ◯ Epidemiology study

Answer

Phase 3 clinical investigation - **This is the correct answer.**
Determining the overall risk versus the benefit of a new pharmaceutical is part of Phase 3 clinical study. The risk versus benefit is one of the last steps in the drug evaluation process.

The type of epidemiology study in which individuals are identified according to exposure and followed to determine subsequent disease risk is known as:

- ◯ Cohort study
- ◯ Case control study
- ◯ Cross-sectional study
- ◯ Ecological study

Answer

Cohort study - **This is the correct answer.**
The type of epidemiology study in which individuals are identified according to exposure and followed to determine subsequent disease risk is known as a cohort study. In a cohort study individuals are selected to be part of the group based on their exposure to a particular substance.

An epidemiological study in which the individuals that make up the test cohort are identified according to past exposures is known as:

- ◯ Case control study
- ◯ Prospective cohort study
- ◯ Retrospective cohort study

Answer

Retrospective cohort study - **This is the correct answer.**
This is known as a retrospective cohort study. As the name implies, retrospective cohorts are identified according to past exposure conditions and the follow-up study proceeds forward in time.

This page titled 2.6.6.5.3: Epidemiology Studies is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **5.3: Epidemiology Studies** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.6: Risk Assessment

Learning Objectives

After completing this lesson, you will be able to:

- Identify the basic steps in the risk assessment process.
- Explain the framework for risk-based decision-making.
- Describe methods for identifying hazards.
- Explain methods for toxicity assessment, including dose-response and exposure.

In this section...

Topics include:

2.6.6.6.1: Risk Assessment

2.6.6.6.2: Hazard Identification

2.6.6.6.3: Dose-Response Assessment

2.6.6.6.4: Exposure Assessment

2.6.6.6.5: Risk Characterization

What We've Covered

This section made the following main points:

- A **hazard** is the capability of a substance to cause an adverse effect.
  A **risk** is the probability that the hazard will occur under specific conditions.
- **Risk assessment** is the process of determining hazard, exposure, and risk.
  Risk management is the process of weighing policy alternatives and deciding on the most appropriate regulatory action.
- There are **four basic steps to risk assessment**:
  1. **Hazard Identification**
     - Identify or develop information suggesting or confirming whether a chemical poses a potential hazard to humans.
     - (Quantitative) Structure Activity, or (Q)SAR methods, including computer models, help consider closely related chemicals as a group or category.
     - Read-across involves estimating what a chemical may be like, including the presence or absence of certain properties or activities, based on one or more other chemicals.
     - Adverse Outcome Pathways (AOPs) involve *in vitro* methods that evaluate changes in normal cellular signaling pathways.
     - Other emerging methods include (Quantitative) *in vitro* to *in vivo* extrapolation, or (Q)IVIVE, Integrated Testing Strategies, and Integrated Approaches to Testing and Assessment (IATA).
  2. **Dose-Response Assessment**
     - Carcinogenic (cancer) risk assessment involves two steps:
       1. Perform qualitative evaluation of all epidemiology studies, animal bioassay data, and biological activity.
       2. Quantitation of the risk for substances classified as definite or probably human carcinogens.
     - Non-carcinogenic risk assessment includes:
       - Acceptable Daily Intake (ADI), which divides the NOAEL by uncertainty/safety factors.

- Reference Dose (RfD), which divides the NOAEL or LOAEL by uncertainty/safety factors.
- Benchmark Dose Method (BMD), which extrapolates data to determine a point of departure (POD) that accounts for study quality.
- Assessments for noncancer toxicity effects, acute or short-term exposures, and occupational exposures.

3. **Exposure Assessment**
   - People are exposed to mixtures of hundreds of chemicals in everyday life.
   - An **exposure pathway** describes the:
     - Route a substance takes from its source to its endpoint.
     - How people can be exposed to the substance.
   - The three steps of exposure assessment are to:
     1. Characterize the point of exposure setting and exposure scenario.
     2. Identify exposure pathways.
     3. Quantify the exposure.
   - Exposure models are commonly used because actual exposure measurements are often not available.
4. **Risk Characterization**
   - This final phase predicts the frequency and severity of effects in exposed populations.
   - Biological and statistical uncertainties are described.
   - For carcinogenic risks, the probability of a person developing cancer over a lifetime is estimated by multiplying the cancer slope factor for the substance by the chronic, 70-year average daily intake.
   - For noncarcinogenic effects, the exposure level is compared with an ADI, RfD, or MRL derived for similar exposure periods.

This page titled 2.6.6.6: Risk Assessment is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.6.6.1: Risk Assessment

Did you know?

For many years, the terminology and methods used in human risk or hazard assessment were inconsistent, which led to confusion among scientists, the public, and others.



(Image Source: iStock Photos, ©)

"Red Book" for Risk Assessment (1983)

In 1983, the (U.S.) National Academy of Sciences (NAS) published **Risk Assessment in the Federal Government: Managing the Process**. Often called the "Red Book" by toxicologists and others, it addressed the standard terminology and concepts for risk assessments.



Figure 2.6.6.6.1.1. Toxicology-based approaches to hazard identification, dose-response assessment, exposure analysis, and characterization of risks were described in the 1983 Red Book
(Image Source: National Academies Press)

Key Terms

The following terms are routinely used in risk assessments:

- **Hazard** — capability of a substance to cause an adverse effect.
- **Risk** — probability that the hazard will occur under specific exposure conditions.
- **Risk assessment** — the process by which hazard, exposure, and risk are determined.
- **Risk management** — the process of weighing policy alternatives and selecting the most appropriate regulatory action based on the results of risk assessment and social, economic, and political concerns.

Risk Assessment Steps

The four basic steps in the risk assessment process as defined by the NAS are:

1. **Hazard identification** — characterization of innate adverse toxic effects of agents.
2. **Dose-response assessment** — characterization of the relation between doses and incidences of adverse effects in exposed populations.
3. **Exposure assessment** — measurement or estimation of the intensity, frequency, and duration of human exposures to agents.
4. **Risk characterization** — estimation of the incidence of health effects under the various conditions of human exposure.

Once risks are characterized in step 4, the process of risk management begins (Figure 2).

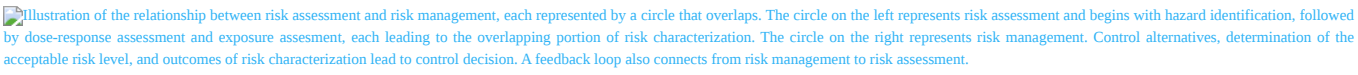


Figure 2.6.6.6.1.2. *Interaction between processes of risk assessment and risk management*
(Image Source: ORAU, ©)

"Silver Book" for Advancing Risk Assessment (2009)

A newer book by the NAS, **Science and Decisions: Advancing Risk Assessment** (2009), often called the “Silver Book” by toxicologists and others, emphasizes uncertainty and variability and cumulative risk, and notes that risk assessment "is at a crossroads."



Figure 2.6.6.6.1.3. *The 2009 Silver Book includes approaches for improving risk analysis and a framework for risk-based decision-making*

(Image Source: National Academies Press)

Risk-Based Decision Making

The co-authors of this Silver Book proposed a framework for risk-based decision-making (Figure 4). The framework consists of three phases:

Enhanced problem formulation and scoping — available risk-management options are identified.



Planning and assessment — risk-assessment tools are used to determine risks under existing conditions and under potential risk-management options.



Risk management — risk and non-risk information is integrated to inform choices among options.



The core of the framework, as noted in the Silver Book, includes the risk assessment paradigm of the Red Book, but differs primarily in its initial and final steps:

- "The framework systematically identifies problems and options that risk assessors should evaluate at the earliest stages of decision-making."
- "It expands the array of impacts assessed beyond individual effects (for example, cancer, respiratory problems, and individual species) to include broader questions of health status and ecosystem protection."
- "It provides a formal process for stakeholder involvement throughout all stages but has time constraints to ensure that decisions are made."
- "It increases understanding of the strengths and limitations of risk assessment by decision-makers at all levels, for example, by making uncertainties and choices more transparent."



Figure 2.6.6.6.1.4. *Framework for risk-based decision-making*
(Source: "Silver Book," chapter 8)

Latest Approaches

Other parts of ToxTutor highlight the latest approaches used in risk assessment. For example:

- **Hazard Identification** describes the emerging approach to hazard identification called **Adverse Outcome Pathways (AOPs)**.
- **Testing for, and Assessing Toxicity** describes the numerous global efforts made, since approximately 1990, to reduce and replace use of laboratory animals in toxicology studies. Included are efforts to develop and validate *in vitro* methods and the emerging use of *in silico* methods.

Knowledge Check

Risk is the:

- ◯ Capability of a substance to cause an adverse effect
- ◯ Weighing of possible alternatives and selecting the most appropriate regulatory actions
- ◯ Probability that a hazard will occur under specific exposure conditions

Answer

Probability that a hazard will occur under specific exposure conditions - **This is the correct answer.**
Risk is the *probability* that a hazard will occur.

In the risk assessment process, what happens during the hazard identification step?

- Characterization of the relation between doses and incidences of adverse effects
- Characterization of innate adverse toxic effects of agents
- Measurement or estimation of intensity, frequency, and duration of human exposures to agents
- Estimation of the incidence of health effects under the various conditions of human exposure

Answer

Characterization of innate adverse toxic effects of agents - **This is the correct answer.**
Hazard identification is the first step in the risk assessment process as defined by the National Academy of Sciences.

What are the phases of the risk-based decision-making framework proposed by the co-authors of the "Silver Book?"

- Enhanced problem formulation and scoping; planning and assessment; and risk management
- Hazard identification; dose-response assessment; exposure assessment; risk characterization
- Hazard identification; risk assessment; action planning

Answer

Enhanced problem formulation and scoping; planning and assessment; and risk management - **This is the correct answer.**
The framework proposed by the co-authors of the "Silver Book" involves enhanced problem formulation and scoping; planning and assessment; and risk management.

This page titled 2.6.6.6.1: Risk Assessment is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **6.1: Risk Assessment** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.6.2: Hazard Identification

The goal of hazard identification in toxicology is to identify or develop information suggesting or confirming that a chemical (or, for example, a consumer product) poses or does not pose a potential hazard to humans.

During earlier years of toxicology, this process relied primarily on human epidemiology data and on various types of animal testing data, supplemented in more recent years with the development of *in vitro* methods such as those focused on assessing the potential for mutations and DNA damage. The future of hazard identification is promising and toxicologists now have various types of *in vitro* methods to explore for hazard identification, along with the emergence of*"chip"*approaches.



Figure 2.6.6.6.2.1. Hazard identification is the first component of risk assessment
(Image Source: ORAU, ©)

These emerging methods are based, in part, on **(Quantitative) Structure Activity, or (Q)SAR methods**. Q(SAR) methods, such as computer models, help toxicologists and others to consider closely related chemicals as a group, or chemical category, rather than as individual chemicals. Not every chemical needs to be tested for every toxicity endpoint, and the data for chemicals and endpoints that have been tested are used to estimate the corresponding properties for other chemicals and endpoints of interest. Data from a chemical category must be judged as adequate to support at least a "screening-level" hazard identification.

One approach involves using endpoint information for one chemical to predict the same endpoint for another chemical that is considered "similar" in some way (such as having structural similarity and similar properties and/or activities).

Read-Across

Another approach for hazard identification used since about 2000 is **read-across**. Read-across can be qualitative or quantitative:

- In **qualitative read-across**, the presence (or absence) of a property/activity such as a particular type of toxic effect for the chemical of interest is inferred from the presence (or absence) of the same property/activity for one or more other chemicals. This qualitative approach provides a "yes/no" answer.
- **Quantitative read-across** uses information for one or more chemicals to estimate what the chemical of interest will be like. Thus, quantitative read-across can be used to obtain a quantitative value for an endpoint, such as a dose-response relationship.

Adverse Outcome Pathways (AOPs)

An emerging approach to hazard identification is the use of **Adverse Outcome Pathways (AOPs)**. AOPs reflect the move in toxicity testing from high-dose studies in laboratory animals to *in vitro* methods that evaluate changes in normal cellular signaling pathways using human-relevant cells or tissues. The AOP concept has emerged as a framework for connecting high throughput toxicity testing (HTT, or high throughput toxicity screening, HTS) and other results.

AOP Learning Channel

The **Human Toxicology Project Consortium** provides a collection of informational videos about AOPs for your further exploration. These videos are available on the **AOP Learning Channel**.

Other Computer Models

Another emerging term is **(quantitative) *in vitro* to *in vivo* extrapolation, or (Q)IVIVE**, used together with what are being called **Integrated Testing Strategies** and Integrated Approaches to Testing and Assessment (IATA).

Toxicology Testing in the 21st Century - A New Strategy

The High Throughput Screening (HTS) Initiative is part of the new toxicology testing strategy developed from the 2004 National Toxicology Program (NTP) **Vision and Roadmap for the 21st Century**.

Traditional toxicological testing is based largely on the use of laboratory animals. However, this approach suffers from low throughput, high cost, and difficulties inherent to inter-species extrapolation – making it of limited use in evaluating the very large number of chemicals with inadequate toxicological data.

NTP recognized that the dramatic technological advances in molecular biology and computer science offered an opportunity to use *in vitro* biochemical- and cell-based assays and non-rodent animal models for toxicological testing. These assays allow for much higher throughput at a much reduced cost. In some assays, many thousands of chemicals can be tested simultaneously in days.

The goal is to move toxicology from a predominantly observational science at the level of disease-specific models to a predominantly predictive science focused upon broad inclusion of target-specific, mechanism-based, biological observations.

The High Throughput Screening program represents a new paradigm in toxicological testing. The HTS program approach to toxicological testing screens for mechanistic targets active within cellular pathways considered critical to adverse health effects such as carcinogenicity, reproductive and developmental toxicity, genotoxicity, neurotoxicity, and immunotoxicity in humans.

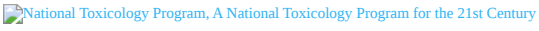


Figure 2.6.6.6.2.2. *National Toxicology Program vision and roadmap*
(Image Source: National Toxicology Program)

Goals of the HTS Program

- To prioritize substances for further in-depth toxicological evaluation.
- To identify mechanisms of action for further investigation (for example, disease-associated pathways).
- To develop predictive models for *in vivo* biological response (predictive toxicology).

Reference:
National Toxicology Program. (2016, March 21). *Tox 21*. U.S. Department of Health and Human Services. Retrieved from **http://ntp.niehs.nih.gov/results/tox21/index.html**

As described in the ***Testing for, and Assessing Toxicity*** section, the EPA is developing "Virtual Tissue Models" such as the Virtual Embryo (v-Embryo™). These types of advanced computer models are being designed to be capable of simulating how chemicals may affect human development and will help reduce dependence on animal study data. They will also provide faster ways of developing chemical risk assessments.

Finally, also noted in the *Testing for and Assessing Toxicity* section, emerging in the toxicologist's tool box are "chip" models (for example, an "organ on a chip"). One example is the "Lung-on-a-chip" that "…offers a new *in vitro* approach to drug screening by mimicking the complicated mechanical and biochemical behaviors of a human lung."



Figure 2.6.6.6.2.3. *Lung-on-a-chip used to mimic pulmonary edema*
(Image Source: The Wyss Institute for Biologically Inspired Engineering)

Knowledge Check

- ◯ Extensive animal testing for toxicity
- ◯ Computer models like (Q)SAR
- ◯ Risk assessment strategies

Answer

Computer models like (Q)SAR - **This is the correct answer.**
Part of the basis for emerging approaches to hazard identification, such as assessing for potential mutations and DNA damage, relies on (Quantitative) Structure Activity (Q)SAR methods.

Adverse Outcome Pathways (AOPs) are methods of hazard identification that:

- ◯ Evaluate changes in normal cellular signaling pathways using human-relevant cells or tissues
- ◯ Identify adverse outcomes in test subjects administered increasing doses of potential toxicants
- ◯ Measure the intensity, frequency, and duration of human exposures to agents
- ◯ Evaluate the health effects under various conditions of human exposure

Answer

Evaluate changes in normal cellular signaling pathways using human-relevant cells or tissues - **This is the correct answer.**
Adverse Outcome Pathways (AOPs) are *in vitro* methods that evaluate changes in normal cellular signaling pathways using human-relevant cells or tissues.

Can quantitative read-across be used to determine the value of an endpoint, such as dose-response relationship?

- ◯ Yes
- ◯ No

Answer

Yes - **This is the correct answer.**
Quantitative read-across can lead to a measurable value for an endpoint, such as a dose-response relationship.

This page titled 2.6.6.6.2: Hazard Identification is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **6.2: Hazard Identification** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.6.3: Dose-Response Assessment

Dose-Response Assessment

The dose-response assessment step of the risk assessment process quantitates the hazards that were identified in the previous step. It determines the relationship between dose and incidence of effects in humans. There are normally two major extrapolations required:

1. From high experimental doses to low environmental doses.
2. From animal doses to human doses.

The procedures used to extrapolate from high to low doses are different for assessing carcinogenic effects and noncarcinogenic effects:

- **Carcinogenic effects** in general are not considered to have a threshold and mathematical models are generally used to provide estimates of carcinogenic risk at very low dose levels.
- **Noncarcinogenic effects** *(for example neurotoxicity)* are considered to have dose thresholds below which the effect does not occur. The lowest dose with an effect in animal or human studies is divided by safety factors to provide a margin of safety.



Figure 2.6.6.6.3.1. Dose-response assessment is a step in the risk assessment process
(Image Source: ORAU, ©)

Carcinogen (Cancer) Risk Assessment

Cancer risk assessment involves two steps:

1. **Perform qualitative evaluation of all epidemiology studies, animal bioassay data, and biological activity *(for example, mutagenicity)*.**The substance is classified as to its carcinogenic risk to humans based on the weight of evidence. If the evidence is sufficient, the substance may be classified as a definite, probable or possible human carcinogen.
2. **Quantitate the risk for those substances classified as definite or probable human carcinogens.** Mathematical models are used to extrapolate from the high experimental doses to the lower environmental doses.

The two primary cancer classification schemes are those of the **Environmental Protection Agency** (EPA) and the **International Agency for Research on Cancer** (IARC). The EPA and IARC classification systems are quite similar.

1. Qualitative Evaluation of Cancer Risk

The EPA's cancer assessment procedures have been used by several Federal and State agencies. The Agency for Toxic Substances and Disease Registry (ATSDR) relies on EPA's carcinogen assessments. A substance is assigned to one of five descriptors shown below in Table 2.6.6.6.3.1.

Descriptor	Definition
Carcinogenic to Humans	Strong evidence of human carcinogenicity
Likely to Be Carcinogenic to Humans	Evidence is adequate to demonstrate carcinogenic potential to humans but does not reach the weight of evidence for the "Carcinogenic to Humans" descriptor.
Suggestive Evidence of Carcinogenic Potential	The weight of evidence is suggestive of carcinogenicity; a concern for potential carcinogenic effects in humans is raised, but the data are judged insufficient for a stronger conclusion.
Inadequate Information to Assess Carcinogenic Potential	Available data are judged inadequate for applying one of the other descriptors. Additional studies generally would be expected to provide further insights.
Not Likely to Be Carcinogenic to Humans	Available data are considered robust for deciding that there is no basis for a substance to be considered a human carcinogen.

Table 2.6.6.6.3.1. *Hazard Descriptors from the EPA's* ***Guidelines for Carcinogen Risk Assessment*** *(March 2005)*

Cancer Data for Humans

The basis for **sufficient human evidence** is an epidemiology study that clearly demonstrates a causal relationship between exposure to the substance and cancer in humans.

The data are determined to be **limited evidence in humans** if there are alternative explanations for the observed effect.

The data are considered to be **inadequate evidence in humans** if no satisfactory epidemiology studies exist.

Cancer Data for Animals

An increase in cancer in more than one species or strain of laboratory animals or in more than one experiment is considered **sufficient evidence in animals**. Data from a single experiment can also be considered sufficient animal evidence if there is a high incidence or unusual type of tumor induced. Normally, however, a carcinogenic response in only one species, strain, or study is considered as only **limited evidence in animals**.

2. Quantitative Evaluation of Cancer Risk

When an agent is classified as a Human or Probable Human Carcinogen, it is then subjected to a **quantitative risk assessment**. For those designated as a Possible Human Carcinogen, the risk assessor can determine on a case-by-case basis whether a quantitative risk assessment is warranted.

The key risk assessment parameter derived from the EPA carcinogen risk assessment is the **cancer slope factor**. This is a toxicity value that quantitatively defines the relationship between dose and response. The cancer slope factor is a plausible upper-bound estimate of the probability that an individual will develop cancer if exposed to a chemical for a lifetime of 70 years. The cancer slope factor is expressed as mg/kg/day.

Linearized Multistage Model (LMS)

Mathematical models are used to extrapolate from animal bioassay or epidemiology data to predict low-dose risk. Most assume linearity with a zero threshold dose.

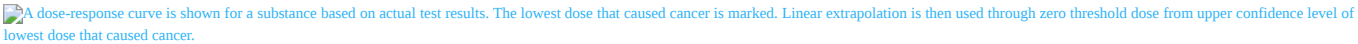


Figure 2.6.6.6.3.2. *The Linearized Multistage Model is used to extrapolate cancer risk from a dose-response curve using the cancer slope factor*
(Image Source: NLM)

EPA uses the **Linearized Multistage Model (LMS)** illustrated in Figure 2 to conduct its cancer risk assessments. It yields a cancer slope factor, known as the **q1*** (pronounced "Q1-star"), which can be used to predict cancer risk at a specific dose. It assumes

linear extrapolation with a zero dose threshold from the upper confidence level of the lowest dose that produced cancer in an animal test or in a human epidemiology study.

Other Models

Other models that have been used for cancer assessments include:

- **One-hit model**, which assumes there is a single stage for cancer and that one molecular event induces a cell transformation. This is a very conservative model.
- **Multi-hit model**, which assumes several interactions are needed before a cell can be transformed. This is one of the least conservative models.
- **Probit model**, which assumes log normal distribution (Probit) for tolerances of exposed population. This model is sometimes used, but generally considered inappropriate for assessing cancer risk.
- **Physiologically Based Pharmacokinetic (PBPK) Models**, which incorporate pharmacokinetic and mechanistic data into the extrapolation process. This model requires extensive data and is becoming commonly used.

Application of Models to Estimate Chemical Concentrations in Drinking Water

The chemical chlordane has been found to cause a lifetime risk of one cancer death in a million persons. Different cancer risk assessment models vary in their estimates of drinking water concentrations for chlordane as illustrated in Table 2.6.6.6.3.2:

Model	Concentration (µg/L)
Probit	50
Multi-hit	2
Linearized multistage	0.07
One-hit	0.03

Table 2.6.6.6.3.2. *Estimates of drinking water chlordane concentrations by various cancer assessment models*

PBPK models are relatively new and are being employed when biological data are available. They quantitate the absorption of a foreign substance, its distribution, metabolism, tissue compartments, and elimination. Some compartments store the chemical (such as bone and adipose tissue) whereas others biotransform or eliminate it (such as liver or kidney). All these biological parameters are used to derive the target dose and comparable human doses.

Noncarcinogenic Risk Assessment

Historically, the **Acceptable Daily Intake (ADI)**procedure has been used to calculate permissible chronic exposure levels for humans based on noncarcinogenic effects. The ADI is the amount of a chemical to which a person can be exposed each day for a long time (usually lifetime) without suffering harmful effects. It is determined by applying safety factors (to account for the uncertainty in the data) to the highest dose in human or animal studies that has been demonstrated not to cause toxicity (NOAEL).

The EPA has slightly modified the ADI approach and calculates a **Reference Dose (RfD)** as the acceptable safety level for chronic noncarcinogenic and developmental effects. Similarly, the ATSDR calculates **Minimal Risk Levels** (MRLs) for noncancer endpoints.

The **critical toxic effect** used in the calculation of an ADI, RfD, or MRL is the serious adverse effect that occurs at the lowest exposure level. It may range from lethality to minor toxic effects. It is assumed that humans are as sensitive as the animal species unless evidence indicates otherwise.

Assessment of Chronic Exposures

In determining the ADIs, RfDs or MRLs, the **NOAEL** is divided by safety factors (uncertainty factors) in order to provide a margin of safety for allowable human exposure.

ADI (human dose) equals NOAEL (experimental dose) divded by safety factors

When a NOAEL is not available, a **LOAEL** can be used to calculate the RfD.

An additional safety factor is included if a LOAEL is used. A Modifying Factor of 0.1–10 allows risk assessors to use scientific judgment in upgrading or downgrading the total uncertainty factor based on the reliability and quality of the data. For example, if a particularly good study is the basis for the risk assessment, a modifying factor of <1 may be used. If a poor study is used, a factor of >1 can be incorporated to compensate for the uncertainty associated with the quality of the study.

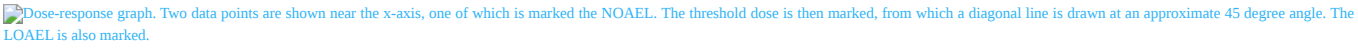


Figure 2.6.6.6.3.3. *Dose-response curve for noncarcinogenic effects*
(Image Source: NLM)

Figure 3 above shows a dose-response curve for noncarcinogenic effects which also identifies the NOAEL and LOAEL. Any toxic effect might be used for the NOAEL/LOAEL so long as it is the most sensitive toxic effect and considered likely to occur in humans.

The **Uncertainty Factors** or **Safety Factors** used to derive an ADI or RfD are listed in Table 2.6.6.6.3.3.

Situation	Uncertainty/Safety Factor
Human variability	10x
Extrapolation from animals to humans	10x
Use of less than chronic data	10x
Use of LOAEL instead of NOAEL	10x
Modifying factor	0.1—10x

Table 2.6.6.6.3.3. Uncertainty/Safety factors used to derive an Acceptable Daily Intake (ADI) or Reference Dose (RfD)

The modifying factor is used only in deriving EPA Reference Doses. The number of factors included in calculating the ADI or RfD depends upon the study used to provide the appropriate NOAEL or LOAEL.

The general formula for deriving the RfD is:



The more uncertain or unreliable the data become, the higher the total uncertainty factor that is applied. An example of an RfD calculation is provided below. A subchronic animal study with a LOAEL of 50 mg/kg/day was used in the numerator. Uncertainty factors used in the denominator are 10 for human variability, 10 for an animal study, 10 for less than chronic exposure, and 10 for use of an LOAEL instead of a NOAEL.



In addition to chronic effects, RfDs can also be derived for other long-term toxic effects, including developmental toxicity.

Traditionally, the **NOAEL method** has been used to determine the **point of departure (POD)** from animal toxicology data for use in risk assessments. However, this approach has limitations such as a strict dependence on the dose selection, dose spacing, and sample size of the study from which the critical effect has been identified. Also, using the NOAEL does not take into consideration the shape of the dose-response curve and other related information.

Benchmark Dose Method

The **benchmark dose (BMD) method**, first proposed as an alternative in the 1980s, addresses many limitations of the NOAEL method. It is less dependent on dose selection and spacing and takes into account the shape of the dose-response curve (Figure 4). In addition, the estimation of a BMD 95% lower bound confidence limit (BMDL) results in a POD that appropriately accounts for study quality (i.e., sample size). With the availability of user-friendly BMD software programs, including the EPA's Benchmark Dose Software (BMDS), the BMD has become the method of choice for many health organizations worldwide.

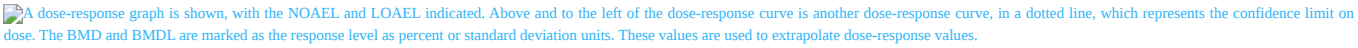


Figure 2.6.6.6.3.4. *Extrapolated values using the benchmark dose method reflect the shape of a dose-response curve (Image Source: EPA)*

Assessment of Noncancer Toxicity Effects

While the Agency for Toxic Substances and Disease Registry (ATSDR) does not conduct cancer risk assessments, it does derive **Minimal Risk Levels (MRLs)**for noncancer toxicity effects (such as birth defects or liver damage). The MRL is defined as an estimate of daily human exposure to a substance that is likely to be without an appreciable risk of adverse effects over a specified duration of exposure. For inhalation or oral routes, MRLs are derived for acute (14 days or less), intermediate (15–364 days), and chronic (365 days or more) durations of exposures.

The method used to derive MRLs is a modification of the EPA's RfD methodology. The primary modification is that the uncertainty factors of 10 may be lower, either 1 or 3, based on scientific judgment. These uncertainty factors are applied for human variability, interspecies variability (extrapolation from animals to humans), and use of a LOAEL instead of NOAEL. As in the case of RfDs, the product of uncertainty factors multiplied together is divided into the NOAEL or LOAEL to derive the MRL.

Assessment of Acute or Short-Term Exposures

Risk assessments are also conducted to derive permissible exposure levels for acute or short-term exposures to chemicals. Health Advisories (HAs) are determined for chemicals in drinking water. HAs are the allowable human exposures for 1 day, 10 days, longer-term, and lifetime durations. The method used to calculate HAs is similar to that for the RfDs using uncertainty factors. Data from toxicity studies with durations of length appropriate to the HA are being developed.

Assessment of Occupational Exposures

For **occupational exposures**, Permissible Exposure Levels (PELs), Threshold Limit Values (TLVs), and National Institute for Occupational Safety and Health (NIOSH) Recommended Exposure Levels (RELs) are developed. They represent dose levels that will not produce adverse health effects from repeated daily exposures in the workplace. The method used to derive is conceptually the same. Safety factors are used to derive the PELs, TLVs, and RELs.

Conversion of Animal Doses to Human Dose Equivalents

Animal doses must be converted to human dose equivalents. The **human dose equivalent** is based on the assumption that different species are equally sensitive to the effects of a substance per unit of body weight or body surface area.

Historically, the FDA used a ratio of body weights of humans to animals to calculate the human dose equivalent. The EPA has used a ratio of surface areas of humans to animals to calculate the human dose equivalent. Some current approaches include multiplying the animal dose by the ratio of human to animal body weight raised to either the 2/3rd or 3/4th power (to convert from body weight to surface area). Toxicologists and risk assessors should check to make sure that the approach they are using is the one mandated or recommended by the regulatory agency of most relevance to their efforts.

Allowable Exposures to Contamination Sources

The last step in risk assessment is to express the risk in terms of allowable exposure to a contaminated source. Risk is expressed in terms of the concentration of the substance in the environment where human contact occurs. For example, the unit for assessing risk in air is risk per -3 g)." tabindex="0">mg/m^3 whereas the unit for assessing risk in drinking water is risk per -3 g)." tabindex="0">mg/L.

For carcinogens, the media risk estimates are calculated by dividing cancer slope factors by 70 kg (average weight of a man) and multiplying by 20 m^3/day (average inhalation rate of an adult) or 2 liters/day (average water consumption rate of an adult).

Knowledge Check

The procedures used to extrapolate from high to low doses primarily depend upon the:

- ◯ Threshold dose of the substance
- ◯ Rate of lethality in laboratory animals

- ◯ Carcinogenicity of the substance

Answer

Genotoxic carcinogenicity of the substance - **This is the correct answer.**
The procedure for extrapolation from high to low doses depend on whether or not the effects are carcinogenic. Carcinogenic effects are not considered to have a threshold dose and mathematical models are used to estimate the risk of carcinogenicity at very low doses. Noncarcinogenic effects are considered to have threshold doses and the margin of safety (MOS) is calculated.

According to EPA, a substance is classified as likely to be carcinogenic to humans when:

- ◯ There is strong evidence of human carcinogenicity
- ◯ Evidence is adequate to demonstrate potential carcinogenicity to humans, but not strongly enough to definitively classify as carcinogenic
- ◯ The weight of evidence suggests human carcinogenicity, but the data are determined not to be sufficient for a stronger conclusion
- ◯ Robust data lead to the conclusion that a substance is clearly carcinogenic to humans

Answer

Evidence is adequate to demonstrate potential carcinogenicity to humans, but not strongly enough to definitively classify as carcinogenic - **This is the correct answer.**
A substance is classified as likely to be carcinogenic to humans when evidence is adequate to demonstrate carcinogenic potential to humans but does not reach the weight of evidence for the descriptor Carcinogenic to Humans.

The primary cancer risk assessment model used by the EPA is known as the:

- ◯ Linearized Multistage Model (LMS)
- ◯ Probit Model
- ◯ Physiologically Based Pharmacokinetic Model (PB-PK)

Answer

Linearized Multistage Model (LMS) - **This is the correct answer.**
EPA uses the Linearized Multistage Model (LMS) to conduct its cancer risk assessments, producing the q1* that is used to predict cancer risk at a specific dose.

The Acceptable Daily Intake (ADI) is calculated by:

- ◯ Dividing the NOAEL by safety factors
- ◯ Dividing the NOAEL by the LOAEL
- ◯ Multiplying the RfD by a modifying factor
- ◯ Linear extrapolation from the LOAEL to the zero intercept

Answer

Dividing the NOAEL by safety factors - **This is the correct answer.**
The ADI is calculated by dividing the NOAEL by safety factors.

Animal doses must be converted to human dose equivalents for risk assessment. When doing this, toxicologists and risk assessors must:

- ◯ Multiply the animal dose by the ratio of human to animal body weight raised to the $^{2}/_{3}$ power

- ◯ Ensure they use the conversion method mandated or recommended by the regulatory agency most relevant to their efforts
- ◯ Multiply the animal dose by the ratio of human to animal body weight raised to the $^{3}/_{4}$ power

Answer

Ensure they use the conversion method mandated or recommended by the regulatory agency most relevant to their efforts - **This is the correct answer.**
Toxicologists and risk assessors should check to ensure they use the approach mandated or recommended by the regulatory agency most relevant to their efforts.

Minimal Risk Levels (MRLs) are derived:

- ◯ Similarly to deriving the RfD, but with a potentially lower uncertainty factor
- ◯ By multiplying the cancer slope factor by the lowest exposure dose
- ◯ By multiplying the LOAEL by safety factors

Answer

Similarly to deriving the RfD, but with a potentially lower uncertainty factor - **This is the correct answer.**
The MRL is calculated much like the RfD, except that the uncertainty factors of 10 may be lower (1 or 3), based on scientific judgment.

This page titled 2.6.6.6.3: Dose-Response Assessment is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **6.3: Dose-Response Assessment** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Exposure Assessment

No Exposure = No Risk

An expression used in toxicology is "no exposure = no risk." Exposure assessment is a key step in the risk assessment process because without an exposure, even the most toxic chemical does not present a threat. Our understanding of potential exposures to chemicals has grown significantly since approximately 1980. For example, research has identified previously "missing" sources and pathways of potential indoor air exposures such as chemicals from consumer products or elsewhere that end up in household dust.

Environmental contaminants are analyzed according to their releases, movement and fate in the environment, and the exposed populations. Consumer products and pharmaceuticals are analyzed in terms of reasonably foreseeable potential exposures.

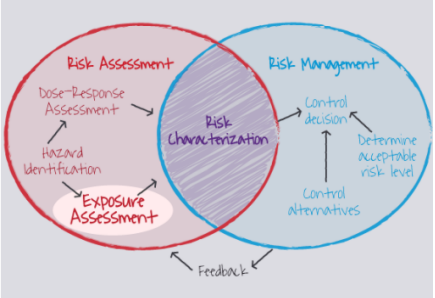


Figure 2.6.6.6.4.1. Exposure assessment as a component of risk assessment
(Image Source: ORAU, ©)

Sources of Exposure

Everyday life can involve a person being exposed to mixtures of hundreds of chemicals. Examples of documented non-occupational sources include:

- Consumer products (perhaps 20 or more during a day).
- Clothing.
- Residential and other water.
- Indoor and outdoor air.
- Food and food packaging.
- Beverages.
- Toys.
- Furniture.
- Carpeting, paint, and other building materials.
- Tobacco smoke.
- Household dust containing chemicals from consumer products.
- Outdoor soil and soil tracked indoors.

Figure 2.6.6.6.4.2 highlights potential exposures from a chemical used in furniture fabric.

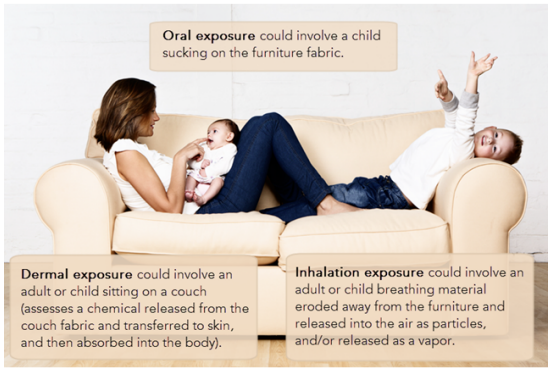


Figure 2.6.6.6.4.2. Exposure to a chemical used in furniture textiles
(Image Source: Adapted from U.S. National Academy Press publication and iStock Photos, ©)

Figure 2.6.6.6.4.3 illustrates some of the types of possible consumer exposures from a hand dishwashing product.



Figure 2.6.6.6.4.3. The many uses and types of exposures for a dishwashing product
(Image Source: Adapted from iStock Photos, ©)

Other residential sources of chemicals can come from household air and water. For example, chemicals in air can deposit on, absorb into, or **adsorb** (Adsorption occurs when molecules of a gas, liquid, or dissolved solid adhere to a surface and create a thin film around it) onto household materials (such as carpets and foods and food packaging), which can lead to dermal and oral exposures. When chemicals are confined to indoor spaces and not diluted in outdoor air, there can be large differences in indoor versus outdoor levels of a chemical.

Figure 2.6.6.6.4.4 illustrates some other residential exposures to chemicals via water.

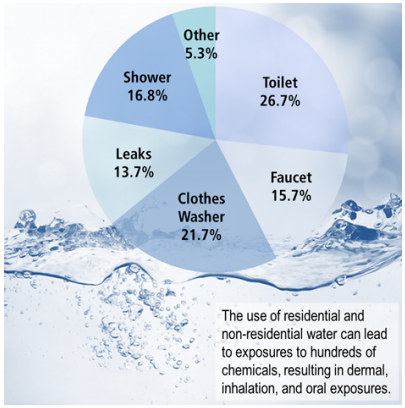


Figure 2.6.6.6.4.4. Water is a household source of chemicals
(Data Source: American Water Works Association Research Foundation, "Residential End Users of Water," 1999.
Learn more at EPA WaterSense.
Image Source: Adapted from iStock Photos, ©)

Exposure Pathways

The route a substance takes from its source (where it began) to its endpoint (where it ends), and how people can come into contact with (or be exposed to) it is defined as an **exposure pathway**.
An exposure pathway has five parts:

1. A **source of exposure**, such as using a consumer product for a household task or a chemical spilled from a truck onto a highway.
2. An environmental media and transport mechanism, such as movement through the indoor or outdoor air or groundwater.
3. A point of exposure, such as a person's house or a private well.
4. A route of exposure — eating, drinking, breathing, or touching.
5. A receptor population — a person or group of people potentially or actually exposed.

When all five parts are present, the exposure pathway is termed a **completed exposure pathway**.
(Source: ATSDR Glossary)

Process of Exposure Assessment

Exposure assessment is a three-step process:



The main variables in the exposure assessment are:

- Exposed populations.
- Types of substances.
- Single substance or mixture of substances.
- Frequency and duration of exposure.
- Pathways and types of exposure.

All possible types of reasonably foreseeable exposures are considered in order to assess the toxicity and risk that might occur.

Considerations for Environmental Exposure

For an environmental exposure, the risk assessor would look at the physical environment and the potentially exposed populations. The physical environment may include considerations about climate, vegetation, soil type, groundwater and surface water. Populations that may be exposed as the result of chemicals that migrate from the site of pollution are also considered.

Subpopulations may be at greater risk due to a higher level of exposure or because they have increased sensitivity. Examples include infants, the elderly, pregnant women, and those with chronic illness.

Pollutants may be transported away from the source and may be physically, chemically, or biologically transformed. They may also accumulate in various materials. Assessment of the chemical fate requires knowledge of many factors, including:

- Organic carbon and water partitioning at equilibrium (Koc).
- Chemical partitioning between soil and water (Kd).
- Partitioning between air and water (Henry's Law Constant).
- Solubility constants.
- Vapor pressures.
- Partitioning between water and octanol (Kow).
- Bioconcentration factors.

These factors are integrated with the data on sources, releases, and routes of the pollutants to determine the exposure pathways of importance, such as groundwater, surface water, air, soil, food, and/or breastmilk.

Use of Exposure Models

Because actual measurements of exposures are often not available, exposure models may be used. For example:

- In air quality studies, chemical emission and air dispersion models are used to predict the air concentrations to downwind residents.
- Residential wells downstream from a site may not currently show signs of contamination. They may become contaminated in the future as chemicals in the groundwater migrate to the well site.

In these situations, groundwater transport models can be used to estimate when chemicals of potential concern will reach the wells.

Information Sources on Chemical Exposures and Health

The future of exposure assessment promises to involve more information and more approaches.
Find out more information about:

- Chemicals in personal care products from the National Institute of Environmental Health Sciences (NIEHS).
- Data connections between the environment and diseases listed in the Comparative Toxicogenomics Database (CTD).
- Chemical safety tools and databases from the Environmental Protection Agency (EPA).
- Health effects of chemical exposures fact sheet and Web site information from the Agency for Toxic Substances & Disease Registry (ATSDR).
- Publications like "Exploring Global Exposure Factors Resources for Use in Consumer Exposure Assessments."

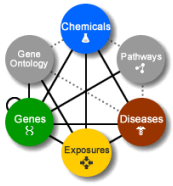


Figure 2.6.6.6.4.5. Topics covered by CTD
(Image Source: CTD)

Considerations When Reading About a New Exposure Study

When reading about a new exposure study, there are several questions you can consider to critically evaluate studies about everyday chemical exposures (source):

1. Is the study published in a peer-reviewed journal?
2. Are there other publications that lend or do not lend support to the current research, or is this study possibly the first of its kind that suggests an emerging issue for regulators, chemical companies, consumer product manufacturers, and others to consider?
3. In addition to scientific journal publications, is there information available from government or other Web sites that provides perspectives about the exposures and potential risks?
4. How was the study conducted? For example, is it a preliminary or "pilot" study involving a small number of people who were studied, and did the participants represent a narrow or broad, wide range of the types of potentially affected consumers?
5. Did the study use household products/materials or food to which some consumers are likely to be exposed? If yes, are there geographical or other limitations that should be noted by the authors such as the products or food being likely to be sold only in one country or region of the world?
6. Did the study try to approach "reasonably foreseeable" consumer exposure conditions?
7. Is there a known or reasonably foreseeable association between these types of exposures and human adverse effects?

This page titled 2.6.6.6.4: Exposure Assessment is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **6.4: Exposure Assessment** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Risk Characterization

Approaches to risk characterization continue to evolve. The final stage in the risk assessment process involves predicting the frequency and severity of effects in exposed populations. The conclusions reached from the stages of hazard identification and exposure assessment are integrated to determine the probability of effects likely to occur in humans exposed under similar conditions.

Because most risk assessments include major uncertainties, it is important to describe biological and statistical uncertainties in risk characterization. The assessment should identify which components of the risk assessment process involve the greatest degree of uncertainty.

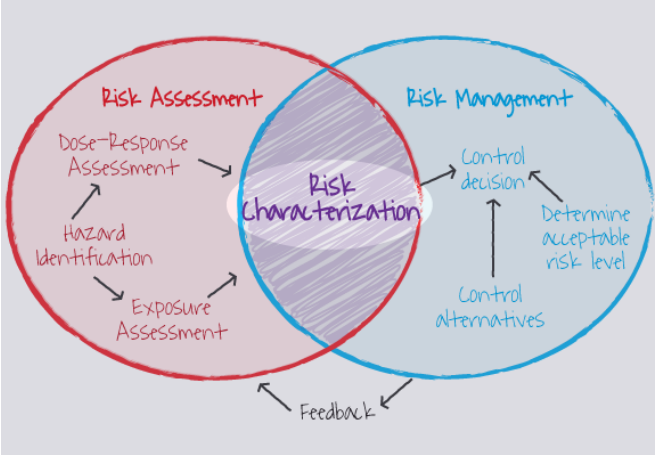


Figure 1. Risk characterization is the final phase of risk assessment (Image Source: ORAU, ©)

For Carcinogenic Risks

Potential human carcinogenic risks associated with chemical exposure are expressed in terms of an increased probability of developing cancer during a person's lifetime. For example, a 10-6 increased cancer risk represents an increased lifetime risk of 1 in 1,000,000 for developing cancer. For carcinogenicity, the probability of an individual developing cancer over a lifetime has historically been estimated by multiplying the cancer slope factor *(mg/kg/day)* for the substance by the chronic *(70-year average)* daily intake *(mg/kg/day)*.

For Noncarcinogenic Effects

For noncarcinogenic effects, the exposure level has historically been compared with an ADI, RfD, or MRL derived for similar exposure periods. Three exposure durations are considered: acute, intermediate, or chronic. For humans:

- **Acute effects** — arise within days to a few weeks.
- **Intermediate effects** — evident in weeks to a year.
- **Chronic effects** — manifest in a year or more.

For Multiple Exposures

In some complex risk assessments, such as for hazardous waste sites, the risk characterization must consider multiple chemical exposures and multiple exposure pathways (described in Exposure Assessments).
Simultaneous exposures to several chemicals, each at a subthreshold level, can often cause adverse effects by "adding" the multiple exposures together, called **dose additivity**.
The assumption of dose additivity is most acceptable when substances induce the same toxic effect by the same mechanism. When available, information on mechanisms of action and chemical interactions are considered and are useful in deriving more scientific risk assessments.
Individuals are often exposed to a substance by more than one exposure pathway*(for example, drinking contaminated water and inhaling contaminated dust)*.

Knowledge Check

1) A major component of exposure assessment involves:

a) Identifying the exposure pathways
b) Measuring the amount of a substance that is metabolized in the body
c) Determining the amount of exposure that must be reduced to comply with the acceptable risk level

Answer

Identifying the exposure pathways - **This is the correct answer.**
Exposure pathways are key to exposure assessment because they identify the route a substance takes from its source to its end point, as well as how people can be exposed to the substance.

2) The movement of substances in environmental media is primarily predicted by:

a) Tagging substances with radioactive tracers and measuring radioactivity at various times and locations within the environmental media
b) Using exposure models to derive scientific estimates
c) Performing actual measurements of exposure pathways

Answer

Using exposure models to derive scientific estimates - **This is the correct answer.**
Since actual measurements of exposures are often unavailable, exposure models may be used.

3) Which of the following is not true about risk characterization?

a) It involves predicting the frequency and severity of effects in exposed populations
b) It determines the amount of exposure that must be reduced to comply with the acceptable risk level
c) It integrates conclusions reached in hazard identification and exposure assessment
d) It yields probabilities of effects likely to occur in humans exposed under similar conditions

Answer

It determines the amount of exposure that must be reduced to comply with the acceptable risk level - **This is the correct answer.**
Risk characterization involves predicting the frequency and severity of effects in exposed populations. It integrates conclusions reached in hazard identification and exposure to yield probabilities of effects likely to occur in humans exposed under similar conditions.

4) An increased cancer risk of 2.0 X 10^-6 (^ indicates to the -6 power) means that:

a) It is likely that two people out of one million will develop the specific type of cancer in their lifetime due to exposure to the chemical
b) The xeniobiotic for which the cancer risk assessment was performed is likely to cause cancer in two people on a yearly basis
c) It is likely that two people out of one thousand will develop the specific type of cancer in their lifetime due to exposure to the chemical
d) It is probable that two million people will develop cancer if they are continuously exposed to the chemical for life

Answer

It is likely that two people out of one million will develop the specific type of cancer in their lifetime due to exposure to the chemical - **This is the correct answer.**
An increased cancer risk of 2 times 10^-6 means two in a million people will likely develop the specific type of cancer in their lifetime due to exposure to the chemical.

This page titled 2.6.6.6.5: Risk Characterization is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **6.5: Risk Characterization** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.7: Exposure Standards and Guidelines

Learning Objectives

After completing this lesson, you will be able to:

- Explain the difference between exposure standards and guidelines.
- Identify approaches to regulating consumer products and drug safety.
- Describe standards and guidelines for environmental and occupational exposure.

In this section...

Topics include:

2.6.6.7.1: Exposure Standards and Guidelines

2.6.6.7.2: Regulation of Consumer Products and Drug Safety

2.6.6.7.3: Environmental Exposure Standards/Guidelines

2.6.6.7.4: Occupational (Workplace) Exposure Standards/Guidelines/Approaches

What We've Covered

This section made the following main points:

- **Standards** are legally acceptable exposure levels or controls set by Congressional or Executive mandate.
- **Guidelines** are recommended maximum exposure levels and are voluntary and not legally enforceable.
- Consumer products
  - The U.S. Consumer Product Safety Commission (CPSC) protects the public from unreasonable risks of harm connected with consumer products.
  - The CPSC establishes consumer exposure standards for hazardous substances and articles.
  - The CPSC requires warning labels on containers of household products that are toxic, corrosive, irritating, or sensitizing.
- Drugs
  - FDA approval is required before pharmaceuticals can be marketed.
  - Animal studies and human clinical trials are required to determine toxic dose levels.
  - The New Drug Application (NDA) contains guidance for drug usage and warnings regarding side effects and interactions.
  - Information about a drug's harmful side effects must be provided through labeling and package inserts, publication in the Physicians' Desk Reference (PDR), and direct-to-consumer marketing.
- Food additives
  - The FDA is responsible for approving food additives.
  - **Direct additives** are intentionally added to foods for functional purposes and include processing aids, flavors, appearance agents, and nutritional supplements.
  - **Indirect additives** are not intentionally added to foods and are not natural constituents of foods, but become constituents during production, processing, packaging, and storage.
  - FDA scientists must review new direct food additives before they can be used in foods.
  - Generally Recognized as Safe (GRAS) additives are generally accepted as safe for an intended use and can be introduced into the food supply without prior FDA approval.
- Environment

- The EPA establishes exposure standards for pesticides, water pollutants, air pollutants, and hazardous wastes.
- Pesticides must be registered with EPA after undergoing extensive analyses.
- The EPA prepares health advisories (HAs) as voluntary exposure guidelines for drinking water contamination.
- Ambient water quality criteria help control pollution sources at the point of release into the environment.
- National Ambient Air Quality Standards (NAAQS) protect public health and welfare from air pollution.
- Hazardous wastes are regulated under the Resource Conservation and Recovery Act (RCRA) and Superfund.
- RCRA regulates hazardous and non-hazardous solid waste.

- Occupational Safety
  - The Occupational Safety and Health Administration (OSHA) establishes legal standards for worker exposure in the United States.
  - Permissible Exposure Limits (PELs) list air concentration limits for chemicals, but not skin absorption or sensitization.
  - Short Term Exposure Limit (STELs) PELs are concentration limits of substances in the air that workers may be exposed to for 15 minutes without adverse effects.
  - Ceiling limits are concentration limits for airborne substances that must not be exceeded.
  - Immediately dangerous to life or health (IDLH) designates an airborne exposure or atmosphere that could lead to death or immediate or delayed permanent adverse health effects.
  - Control banding (CB) determines a control measure based on a band of hazards, such as skin irritation or carcinogenic potential, and exposures.

This page titled 2.6.6.7: Exposure Standards and Guidelines is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

Exposure Standards and Guidelines

Exposure standards and guidelines are developed by governments to protect the public from harmful substances and activities that can cause serious health problems. This section describes standards and guidelines relating to protection from the toxic effects of chemicals only.

Exposure standards and guidelines are determined by risk management decisions. Risk assessments provide regulatory agencies with estimates of numbers of persons potentially harmed under specific exposure conditions. Regulatory agencies then propose exposure standards and guidelines designed to protect the public from "unacceptable risk" levels.

Exposure standards and guidelines usually provide numerical exposure levels for various media (such as food, consumer products, water, and air) that should not be exceeded. Alternatively, these standards may be preventive measures to reduce exposure (such as labeling, special ventilation, protective clothing and equipment, and medical monitoring).



Figure 2.6.6.7.1.1. Standards and guidelines protect the public from harmful substances
(Image Source: iStock Photos, ©)

Standards and Guidelines

More specifically, standards and guidelines for chemical exposure levels consist of the following:

- **Standards** — legally acceptable exposure levels or controls issued as the result of Congressional or Executive mandate. They result from formal rulemaking and are legally enforceable. Violators are subject to punishment, including fines and imprisonment.
- **Guidelines** — recommended maximum exposure levels which are voluntary and not legally enforceable. Guidelines may be developed by regulatory and non-regulatory agencies, or by some professional societies.

Federal and state regulatory agencies have the authority to issue permissible exposure standards and guidelines in the following categories:

- Consumer Product Exposure Standards and Guidelines
- Environmental Exposure Standards and Guidelines
- Occupational Exposure Standards and Guidelines

Knowledge Check

1) Exposure standards are:

Developed by chemical manufacturers
Recommended maximum exposure levels which are voluntary and not legally enforceable
Legally enforceable acceptable exposure levels or controls

Answer

Legally enforceable acceptable exposure levels or controls - **This is the correct answer.**
Exposure standards are legally enforceable acceptable exposure levels or controls resulting from Congressional or Executive mandate.

2) Exposure guidelines are:

Developed by chemical manufacturers
Recommended maximum exposure levels which are voluntary and not legally enforceable
Legally enforceable acceptable exposure levels or controls

Answer

Recommended maximum exposure levels which are voluntary and not legally enforceable - **This is the correct answer.**
Exposure guidelines are recommended maximum exposure levels that are voluntary and not legally enforceable.

This page titled 2.6.6.7.1: Exposure Standards and Guidelines is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **7.1: Exposure Standards and Guidelines** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.7.2: Regulation of Consumer Products and Drug Safety

Regulation of Consumer Products and Drug Safety

Consumer products are often called household products. It is important to know what a category of product is called in the area of the world of interest.

For example, in the United States, cosmetic products are defined by the FDA as those products "intended to be applied to the human body for cleansing, beautifying, promoting attractiveness, or altering the appearance without affecting the body's structure or functions." While cosmetics are often thought of as being make-up products such as eye-liner, lipstick, or nail polish, the FDA definition includes skin-care creams, lotions, powders, and perfumes. However, soap products are excluded from FDA's definition of cosmetics.

In comparison, in the European Union (EU), the European Commission's definition of cosmetic products includes soap, shampoo, deodorant, toothpaste, perfumes, and makeup.

It is also important to keep in mind that some products are known by different names. For example, what is known as a cloth or disposable diaper in the United States is called a nappy in several other countries.

Figure 2.6.6.7.2.1. Shopping
(Image Source: iStock Phot

U.S. Consumer Product Safety Commission (CPSC)

The U.S. Consumer Product Safety Commission (CPSC) is charged with "protecting the public from unreasonable risks of injury or death associated with the use of the consumer products" under its jurisdiction such as toys, cribs, power tools, cigarette lighters, and household chemical-containing products. The CPSC considers if a product could pose a fire, electrical, chemical, or mechanical (such as choking) hazard. CPSC's work, including research, product recalls, education, and administration of regulations, laws, and standards, has resulted in a decline in the rate of deaths and injuries associated with consumer products over the past several decades.
Consumer exposure standards are developed for hazardous substances and articles by the CPSC. The authority under the Federal Hazardous Substance Act (FHSA) pertains to substances other than pesticides, drugs, foods, cosmetics, fuels, and radioactive materials. The CPSC-required warning labels on containers of household products that are toxic, corrosive, irritants, or sensitizers help consumers to safely store and use those products and to give them information about immediate first aid steps to take if an accident happens. Highly toxic substances are labeled with DANGER; less toxic substances are labeled with WARNING or CAUTION.

Figure 2.6.6.7.2.2. CPSC danger label for a gas-powered generator
(Image Source: CPSC)

The CPSC's basis for a determination of highly toxic has been death in laboratory rats at an oral dose of 50 mgs, or an inhaled dose in rats of 200 ppm for one hour, or a 24-hour dermal dose in rabbits of 200 mg/kg. A substance is corrosive if it causes visible destruction or irreversible damage to the skin or eye. If it causes damage that is reversible within 24 hours, it is designated an irritant. An immune response from a standard sensitization test in animals is sufficient for designating the substance a sensitizer.
CPSC is a member of the U.S. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), a permanent committee of the National Institutes of Environmental Health Science (NIEHS) under the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). ICCVAM was created "to establish, wherever feasible, guidelines, recommendations, and regulations that promote the regulatory acceptance of new or revised scientifically valid toxicological tests that protect human and animal health and the environment while reducing, refining, or replacing animal tests and ensuring human safety and product effectiveness."

European Commission General Product Safety Directive (GPSD)

In the EU, the European Commission's General Product Safety Directive (GPSD) requires producers and distributors to place only safe consumer products on the market and to take all necessary measures to prevent risks to consumers. The GPSD excludes certain product categories covered by specific European safety regulations. It provides an alert system (Rapid Alert System for non-food dangerous products – "RAPEX") between the Commission and EU Member States, Norway, Iceland and Liechtenstein.

U.S. Food and Drug Administration (FDA)

The FDA oversees food safety, tobacco products, dietary supplements, prescription and non-prescription drugs, vaccines, blood products, medical devices, electromagnetic radiation emitting devices, cosmetics, animal foods, and veterinary products.

Drugs

Manufacturers of new pharmaceuticals must obtain formal FDA approval before their products can be marketed. Drugs intended for use in humans must undergo animal studies and human clinical trials to determine toxic dose levels prior to filing a New Drug Application (NDA).

Did you know?

During the late 1950s the drug Thalidomide was used to control nausea and vomiting in pregnancy in Canada, Europe, Australia, and parts of Asia. After the drug came on the market, reports appeared of children born with missing limbs, with upper limbs usually more affected than lower limbs. In addition to damage to arms and legs, faces, eyes, ears, and genitalia, internal organs including the heart, kidney, and gastrointestinal tract were damaged.

Thalidomide did not reach the US market due to the efforts of an astute FDA drug reviewer, Dr. Frances Oldham Kelsey, who insisted that thalidomide be fully tested before approval. In response to the episode, the Kefauver Harris Amendment to the Federal Food, Drug, and Cosmetic Act became law in 1962. The amendment requires drug manufacturers to give proof of the effectiveness and safety of a drug before it can be approved.

Thalidomide continues to be prescribed under strict supervision as a treatment for multiple myeloma, leprosy, certain complications of human immunodeficiency virus (HIV), and autoimmune disorders. In the 2000s, children with thalidomide-related birth defects were noted in Brazil due to the use of thalidomide to treat leprosy. Pregnant women were being exposed when family members took the drug.

Figure 2.6.6.7.2.3. Thalido
(Image Source: NLM Cir

An NDA covers all aspects of a drug's effectiveness and safety, including:

- Pharmacokinetics and pharmacological effects.
- Metabolism and postulated mechanism of action.
- Associated risks of the drug.
- Intended uses and therapeutic efficacy.
- Risk:benefit relationship.
- Basis for package inserts supplied to physicians.

The FDA does not issue exposure standards for drugs. Instead, it approves an NDA that contains guidance for usage and warnings concerning potential side effects of a drug. The manufacturer is required to provide this information to physicians prescribing the drug as well as to others that may purchase or use the drug. Information on a drug's harmful side effects is provided in three main ways:

1. Labeling and package inserts that accompany a drug and explain approved uses, recommended dosages, and effects of overexposure.
2. Publication of information in the Physicians' Desk Reference (PDR) and other publically available databases.
3. Direct-to-consumer advertisements.

Figure 2.6.6.7.2.4. Drugs m
(Image Source: iStock Pho

The package insert labels and the PDR contain the following information:

- Description
- Clinical pharmacology
- Indications and usage
- Contraindications
- Warnings
- Precautions
- Adverse reactions
- Interactions
- Overdosage
- Available forms
- Dosage and administration

Figure 2.6.6.7.2.5. Sample package insert for a fictional prescription drug
(Image Source: FDA)

Food Additives

The FDA is responsible for the approval of food additives. Standards are different depending on whether they are direct food additives or indirect food additives.

Direct food additives are intentionally added to foods for functional purposes. Examples of direct food additives include:

- Processing aids
- Texturing agents
- Preservatives
- Flavoring and appearance agents
- Nutritional supplements

Approval usually designates the maximum allowable concentrations in a food product.

Figure 2.6.6.7.2.6. A worker in a food production plant adds a mix of dry flavoring and processing agents to meat
(Image Source: iStock Photos, ©)

Indirect food additives are not intentionally added to foods and they are not natural constituents of foods. They become a constituent of the food product from environmental contamination during production, processing, packaging, and storage. Examples of indirect food additives are:

- Antibiotics administered to cattle.
- Pesticide residues remaining after production or processing of foods.
- Chemicals that migrate from packaging materials into foods.

Exposure standards indicate the maximum allowable concentration of these substances in food.

New direct food additives must undergo stringent review by FDA scientists before they can be approved for use in foods. The manufacturer of a direct food additive must provide evidence of the safety of the food additive in accordance with specified uses. The safety evaluation is conducted by the toxicity testing and risk assessment procedures previously discussed with derivation of the ADI. In contrast to pharmaceutical testing, virtually all toxicity evaluations are conducted with experimental laboratory animals.

FDA approval of all new food additives has been requires since the Food, Drug and Cosmetic Act (FDCA) was amended in 1958. At that time, all existing food additives were Generally Recognized as Safe (GRAS) and no exposure standard was developed. Many of GRAS substances have since been reevaluated and maximum acceptable levels have been established. However, under the law, a substance may be determined to be GRAS for an intended use and introduced into the food supply as such without prior approval by FDA. FDA maintains a searchable database of GRAS substances.

Figure 2.6.6.7.2.7. Select Committee on GRAS Substances searchable database

The FDA reevaluation of GRAS substances requires that specific toxicity tests be conducted based on the level of the GRAS substance in a food product. For example, the lowest level of concern is for an additive used at 0.05 ppm in the food product. Only short-term tests (a few weeks) are required for those compounds. In contrast, a food additive used at levels higher than 1.0 ppm must be tested for carcinogenicity, chronic toxicity, reproductive toxicity, developmental toxicity, and mutagenicity.
The 1958 amendment to the Food, Drug and Cosmetic Act law is known as the Delaney Clause. This clause prohibited the addition of any substance to food that has been shown to induce cancer in man or animals. The implication was that any positive result in an animal test, regardless of dose level or mechanism, is sufficient to prohibit use of the substance. In this case, the allowable exposure level is zero. In 1958, chemical levels could only be measured in parts per thousand whereas analytical methods today allow some chemical levels to be measured down to parts per trillion or quintillion. Such levels might represent negligible cancer risks and Congress has repeatedly amended the Delaney Clause to create more and more exceptions. In 1996, the Delaney Clause was repealed, and the "zero–risk" standard was changed to one of "reasonable certainty."

Food Safety in the European Union (EU)

General Food Law

The basic principles for the EU's food safety policy are defined in the EU's General Food Law (Regulation (EC) No 178/2002), adopted in 2002. This regulation:

- Lays down general principles, requirements and procedures that underpin decision making in matters of food and feed safety, covering all stages of food and feed production and distribution.
- Sets up an independent agency responsible for scientific advice and support, the **European Food Safety Authority (EFSA)** - see below for more information.
- Creates the main procedures and tools for the management of emergencies and crises as well as the Rapid Alert System for Food and Feed (RASFF).

Figure 8. Flags of the European Union and various member nations
(Image Source: iStock Photos, ©

European Food Safety Authority (EFSA)

The European Food Safety Authority (EFSA) was set up in 2002 and is based in Parma, in Italy. It carries out risk assessments before certain foods are allowed to be sold in the EU. The agency was legally established by the EU under the General Food Law - Regulation 178/2002. The General Food Law created a European food safety system in which responsibility for risk assessment (science) and for risk management (policy) are kept separate. EFSA is responsible for the former area, and also has a duty to communicate its scientific findings to the public.
EFSA provides scientific advice to the European Commission and EU countries, to help them take effective decisions to protect consumers. It also plays an essential role in helping the EU respond swiftly to food safety crises. EFSA's remit covers:

- Food and feed safety
- Nutrition
- Animal health and welfare
- Plant protection
- Plant health

It also considers, through environmental risk assessments, the possible impact of the food chain on the biodiversity of plant and animal habitats (discover EFSA topics, EFSA scientific work areas, and application resources by food sector area). EFSA publishes all its scientific outputs, including its scientific opinions, in the EFSA Journal. It also issues a range of supporting publications.
Most of EFSA's work is undertaken in response to requests for scientific advice from the European Commission, the European Parliament and EU Member States. EFSA also carry out scientific work on own initiative, in particular to examine emerging issues and new hazards and to update our assessment methods and approaches. This is known as "self-tasking." EFSA's quality management system (QMS) has been awarded an ISO 9001:2015 certificate, the international standard for quality management.
EFSA's Scientific Panels of experts are responsible for the bulk of EFSA's scientific assessment work. Each of the 10 Panels is dedicated to a different area of the food and feed chain. The Scientific Committee has the task of supporting the work of the Panels on cross-cutting scientific issues. It focuses on developing harmonized risk assessment methodologies in fields where EU-wide approaches are not yet defined.
EFSA staff support the Scientific Panels and Scientific Committee in carrying out most of the Authority's scientific work. The membership of EFSA's Scientific Committee and Panels is renewed every three years (see also Working practices and EFSA Strategy 2020 - Trusted science for safe food).

The EU General Food Law deals with a wide range of issues related to food in general and food safety in particular, including food information and animal welfare. It covers all parts of the food chain from animal feed and food production to processing, storage, transport, import and export, as well as retail sales. It also establishes the principles for risk analysis. These stipulate how, when, and by whom scientific and technical assessments should be carried out in order to ensure that humans, animals, and the environment are properly protected.

EU Food Safety Policy

The EU's food safety policy covers food from farm to fork. The EU food policy comprises:

- Comprehensive legislation on food and animal feed safety and food hygiene.
- Sound scientific advice on which to base decisions.
- Enforcement and checks.

Where specific consumer protection is justified, there may be special rules on:

- Use of pesticides, food supplements, colorings, antibiotics, or hormones.
- Food additives, such as preservatives and flavorings
- Substances in contact with foodstuffs, for example, plastic packaging.
- Labeling of ingredients that may cause allergies.
- Health claims such as "low-fat" or "high-fiber."

The EU's Rapid Alert System for Food and Feed (RASFF) was launched in 1979 and allows information on food and feed to be shared quickly and efficiently between all the relevant bodies at national and EU-level. In a similar vein, the EU Notification System for Plant Health Interceptions (EUROPHYT) is the EU's notification and rapid alert system for plant products entering and being traded within the EU. It helps to prevent the introduction and spread of plant disease and plant pests.
The EU's Trade Control and Expert System (TRACES) is a system for tracking live animals and food and feed of animal origin as they enter the EU and are traded within the EU. It links veterinary authorities across and outside the EU, and enables veterinary services and businesses to react swiftly when a health threat is discovered.

Regulatory Science in the European Union

The following tables describe regulatory science in the EU, including links to the relevant agencies. The list is a work in progress.

Regulatory Science (the table below is best viewed on a computer browser)

Area	Regulatory Science (click links for more information)
Animal health and welfare	European Food Safety Authority
Animal feed	European Food Safety Authority
Biological hazards	European Food Safety Authority
Biocidal Active Substances	European Chemicals Agency
Biocidal Products	European Chemicals Agency
Botanicals	European Food Safety Authority
Consumer products	European Commission General Product Safety Directive (GPSD)
Communicable diseases	European Centre for Disease Prevention and Control
Food additives	European Food Safety Authority
Food ingredients	European Food Safety Authority
Food Colors	European Food Safety Authority
Food Contact Materials	European Food Safety Authority
Food enzymes	European Food Safety Authority
Food Supplements	European Food Safety Authority
Non-plastic food contact materials	European Food Safety Authority
Plastics and plastic recycling	European Food Safety Authority
Smoke Flavorings	European Food Safety Authority
Sweeteners	European Food Safety Authority
Food packaging	European Food Safety Authority
GMO	European Food Safety Authority
Plant Health	European Food Safety Authority
Pesticides	European Food Safety Authority
Nutrition	European Food Safety Authority
Nutrient sources	European Food Safety Authority
Vitamins and minerals	European Food Safety Authority

List above shows regulatory authorities over specific areas of science in the European Union

Regulatory Authorities (the table below is best viewed on a computer browser)

Agency (click links for more information)	Area of Authority
European Food Safety Authority (EFSA)	• Animal health and welfare • Biological hazards • Chemical contaminants • Cross-cutting science • Food ingredients and food packaging • GMO • Nutrition • Pesticides • Plant health • Chemicals • Biocides
European Chemicals Agency (ECHA)	
European Medicines Agency (EMEA)	Scientific evaluation, supervision and safety monitoring of medicines
European Centre for Disease Prevention and Control (ECDC)	Communicable diseases

List above shows regulatory authorities and their specific areas of science in the European Union

Cross-Cutting Science (the table below is best viewed on a computer browser)

Area	Regulatory Agency (click links for more information)
Chemical mixtures	European Food Safety Authority
Cloning	European Food Safety Authority
Emerging risks	European Food Safety Authority
Endocrine active substances	European Food Safety Authority
Margin of Exposure	European Food Safety Authority
Nanotechnology	European Food Safety Authority
Threshold of Toxicological Concern	European Food Safety Authority
Bee health	European Food Safety Authority
Machine learning	European Food Safety Authority

List above shows regulatory authorities over cross-cutting science in the European Union.

Knowledge Check

1) Consumer exposure standards are developed for hazardous substances and articles by the:
a) U.S. Food and Drug Administration (FDA)
b) U.S. Consumer Product Safety Commission (CPSC)
c) General Product Safety Directive (GPSD)

Answer

U.S. Consumer Product Safety Commission (CPSC)

2) Exposure standards for pharmaceuticals are:
a) Issued by the U.S. Food and Drug Administration (FDA)
b) Developed by the Environmental Protection Agency
c) Recommended guidance developed by the U.S. Food and Drug Administration (FDA)

Answer

Recommended guidance developed by the U.S. Food and Drug Administration (FDA)

3) The FDA develops exposure standards for both direct and indirect food additives. Which of the following is an example of an indirect food additive?
a) A preservative added to food products
b) An antibiotic administered to cattle
c) Natural and artificial flavorings
d) A nutritional supplement, such as Vitamin A

Answer

An antibiotic administered to cattle

4) Under the Delaney Clause of 1958, the FDA:

a) Required physicians to strictly adhere to exposure standards for pharmaceuticals
b) Prohibited the addition of any substance to food that has been shown to induce cancer in humans or animals
c) Authorized the addition of potentially carcinogenic substances to food as long as the concentration is at 0.05 ppm or less

Answer

Prohibited the addition of any substance to food that has been shown to induce cancer in humans or animals

5) In the European Union, what regulatory authority is responsible for chemicals and biocides?

a) European Centre for Disease Prevention and Control (ECDC)
b) European Food Safety Authority (EFSA)
c) European Chemicals Agency (ECHA)

Answer

European Food Safety Authority (EFSA)

This page titled 2.6.6.7.2: Regulation of Consumer Products and Drug Safety is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **7.2: Regulation of Consumer Products and Drug Safety** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Environmental Exposure Standards/Guidelines

The Environmental Protection Agency (EPA) is responsible for US-wide laws that require determination and enforcement of environmental exposure standards. In addition, EPA has the authority to prepare recommended exposure guidelines for selected environmental pollutants. Similar organizations exist in US states, and in other countries and groups of countries (such as the European Commission for European Union member countries).
The EPA is responsible for developing exposure standards for:

- Pesticides
- Water pollutants
- Air pollutants
- Hazardous wastes

Pesticides

Pesticides cannot be marketed until they have been registered by the EPA in accordance with the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). In order to obtain registration, a pesticide must undergo an extensive battery of toxicity tests, chemistry analyses, and environmental fate tests.

The Food Quality Protection Act of 1996 (FQPA) changed the way that pesticides were regulated by requiring investigation of:

- Nonoccupational exposure to pesticides.
- The cumulative effects of pesticides having a common toxicity mechanism.
- Any increased susceptibility in infants, children and other sensitive groups.
- Whether the pesticide has endocrine-disrupting effects.

Certain pesticides have been determined by the EPA to pose minimal risk to health or to the environment and are exempt from FIFRA registration.
A primary exposure standard for pesticides is the pesticide tolerance for food use This standard specifies the pesticide residue allowed to remain in or on each treated food product. The EPA residues risk assessment covers:

- Toxicity.
- Amount used and how often.
- How much residue remains by the time a product reaches the market.
- Other ways of being exposed to the pesticide, if any.

Figure 2.6.6.7.3.1. Dusting crops with pest
(Image Source: iStock Photos, ©)

Figure 2.6.6.7.3.2. Aerosol spray cans of p
(Image Source: iStock Photos, ©)

Water Pollutants

Access to safe drinking water and control of water pollution are regulated by the Safe Drinking Water Act (SDWA) and the Clean Water Act (CWA). Under the SDWA, the EPA conducts risk assessments and issues Maximum Contaminant Levels **(MCLs)** for naturally-occurring and man-made contaminants in drinking water. The MCL is the acceptable exposure level which, if exceeded, requires immediate water treatment to reduce the contaminant level.

In addition to establishing MCLs, the EPA can propose recommended exposure guidance for drinking water contamination. As an interim procedure, **maximum contaminant level goals(MCLGs)**may be recommended for long-term exposures to contaminants in drinking water. Generally, no allowable exposure can be recommended for a carcinogenic chemical. When the MCL is issued, an acceptable exposure level based on a cancer risk assessment may be proposed for the MCL.

EPA prepares health advisories **(HAs)** as voluntary exposure guidelines for drinking water contamination. The HAs provide exposure limits for 1-day, 10-day, longer-term, or lifetime exposure periods. They pertain to both cancer and noncancer risks. The formula used to derive a health advisory differs from that for the ADI or RfD in that the HAs pertain to short-term as well as long-term exposures. In addition, human body weight and drinking water consumption are included in the formula.
Data from toxicology studies such as the duration of exposure and the exposure route (such as oral) must be represented in the HAs. For example, the 10-day HA must be based on a NOAEL or LOAEL derivation that was obtained from an animal toxicology study of approximately 10 days duration (routinely 7- to 14-day toxicity studies).
A longer-term HA applies to humans drinking contaminated water for up to 7 years (which could represent 10% of a human's 70-year lifespan). Because 90 days is about 10% of a rat's expected lifespan, the 90-day subchronic study with rats is considered appropriate for providing the basis for the longer-term HA assessment.
A life-time HA (representing lifetime exposure to a toxicant in drinking water) is also determined for noncarcinogens. The procedure uses the RfD risk assessment with adjustments for body weight of an adult human (70 kg) and drinking water consumption of 2 L/day.

Figure 2.6.6.7.3.3. EPA issues health advisories as voluntary exposure guidelines for drinking water contamination
(Image Source: Adapted from EPA)

In addition to drinking water standards, the EPA is authorized under the Clean Water Act (CWA) to issue exposure guidance for controlling pollution in ground water. The intent is to provide clean water for fishing and swimming rather than for drinking purposes. It provides a scheme for controlling the introduction of pollutants into navigable surface water. The recommendations for ground water protection are known as ambient water quality criteria.

The **ambient water quality criteria** are intended to control pollution sources at the point of release into the environment. While these criteria may be less restrictive than the drinking water standards, they usually are the same numeric value. For example, the MCL for drinking water and the ambient water quality criteria for ground water for lead are the same: 0.05 mg/L of water.

Figure 2.6.6.7.3.4. A scientist tests the water for toxicity
(Image Source: iStock Photos, ©)

Air Pollutants

Air emission standards are issued by EPA under the Clean Air Act (CAA). The CAA authorizes the issuance of **National Ambient Air Quality Standards (NAAQS)** for air pollution. There are two types of NAAQS:

1. **Primary NAAQS standards** set limits to protect public health, including people at increased risk (preexisting heart or lung disease, children, older adults)

1. **Primary NAAQS standards** set limits to protect public health, including people at increased risk (preexisting heart or lung disease, children, older adults).
2. **Secondary NAAQS standards** relate to public welfare, such as crops, animals, and structures.

NAAQS have been established for the following major atmospheric pollutants:

- Carbon monoxide
- Sulfur oxide
- Oxides of nitrogen
- Ozone
- Hydrocarbons
- Particulates
- Lead

When air emissions exceed the NAAQS levels, the polluting industry must take action to reduce emissions to acceptable levels.

Figure 2.6.6.7.3.5. A factory releases pollutants into the air
(Image Source: iStock Photos, ©)

Hazardous Wastes

Hazardous wastes are regulated under the Resource Conservation and Recovery Act (RCRA) and the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA), commonly known as **Superfund**. RCRA regulates hazardous and nonhazardous solid waste, including chemical waste produced by industrial processes, medical waste, and underground storage tanks.

The main purpose of CERCLA is to clean up hazardous waste disposal sites. EPA has established standards known as Reportable Quantities **(RQs)**. Companies must report to EPA any chemical release that exceeds the RQ. RQs evaluate physical, chemical, and toxicological properties of a substance. These are called **primary criteria** and include aquatic toxicity, acute mammalian toxicity (oral, dermal, and inhalation), ignitability, reactivity, chronic toxicity, and potential carcinogenicity. **Secondary criteria** evaluate how a substance degrades in the environment.

Minimal Risk Levels (MRLs) for noncancer toxic effects are derived by the Agency for Toxic Substances and Disease Registry (ATSDR), which has a congressional mandate to investigate the health effects of hazardous substances in the environment. MRLs are estimates of daily human exposures that are likely to be without an appreciable risk of adverse effects over a specified duration of exposure. MRLs are derived for acute (14 days or less), intermediate (15–364 days), and chronic (365 days or more) exposures for inhalation or oral routes.

Knowledge Check

1) The exposure standard established by the EPA for the pesticide residue allowed to remain in or on each treated food product is called the:
a) Pesticide tolerance for food use
b) Reportable quantity of pesticides
c) Maximum Contaminant Level (MCL)

Answer

Pesticide tolerance for food use - **This is the correct answer.**
The pesticide tolerance for food use standard established by the EPA specifies the pesticide residue allowed to remain in or on treated food products.

2) The EPA establishes exposure standards for natural and man-made contaminants in drinking water. These standards are called:
a) Ambient Water Quality Criteria
b) Maximum Contaminant Level Goals (MCLGs)
c) Maximum Contaminant Levels (MCLs)

Answer

Maximum Contaminant Levels (MCLs)- **This is the correct answer.**
The EPA conducts risk assessments and issues Maximum Contaminant Levels (MCLs) for naturally-occurring and man-made contaminants in drinking water.

3) What is the difference between primary National Ambient Air Quality Standards (NAAQS) and secondary NAAQS?
a) Primary NAAQS relate to public welfare (for example, crops, animals, and structures); secondary NAAQS set limits to protect public health
b) Primary NAAQS set limits to protect public health; secondary NAAQS relate to public welfare
c) Primary NAAQS are legally enforceable; secondary NAAQS are not

Answer

Primary NAAQS set limits to protect public health; secondary NAAQS relate to public welfare - **This is the correct answer.**
Primary NAAQS set limits to protect public health. Secondary NAAQS relate to public welfare, such as crops, animals, and structures.

4) The Agency for Toxic Substances and Disease Registry (ATSDR) estimates levels for daily human exposure to chemicals that are likely to be without an appreciable risk of adverse effects for specified periods of exposure. These are known as:
a) Minimal Risk Levels (MRLs)
b) Maximum Contaminant Levels (MCLs)
c) Reportable Quantities (RQs)

Answer

Minimal Risk Levels (MRLs) - **This is the correct answer.**
Minimal Risk Levels (MRLs) for noncancer toxic effects estimate the daily human exposures that are likely to be without an appreciable risk of adverse effects over a specific duration of exposure.

This page titled 2.6.6.7.3: Environmental Exposure Standards/Guidelines is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **7.3: Environmental Exposure Standards/Guidelines** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.7.4: Occupational (Workplace) Exposure Standards/Guidelines/Approaches

Occupational Safety and Health Administration (OSHA) Standards

Recommended or mandatory **occupational exposure limits (OELs)** for chemicals exist in many countries. For example, legal standards in the United States are established by the Occupational Safety and Health Administration (OSHA). These standards are known as Permissible Exposure Limits (PELs). The majority of PELs were issued after the 1970 Occupational Safety and Health (OSH) Act.

OSHA maintains the "Permissible Exposure Limits – Annotated Tables" that contain comparative information taken from federal, state, and professional organizations such as the:

- California Division of Occupational Safety and Health (Cal/OSHA) PELs.
- ACGIH® (formerly known as the American Conference of Governmental Industrial Hygienists) Threshold Limit Values (TLVs®).
- National Institute for Occupational Safety and Health (NIOSH) Recommended Exposure Limits (RELs).

These tables list air concentration limits for chemicals but do not include notations for skin absorption or sensitization. The PEL Annotated Tables include the following: Table Z-1, Table Z-2, and Table Z-3.

Figure 2.6.6.7.4.1. Screenshot of a portion of OSHA PEL Annotated Table Z-2

Most OSHA PELs are for airborne substances with allowable exposure limits averaged over an 8-hour day in a 40-hour week. This is known as the **Time-Weighted-Average (TWA) PEL**. While adverse effects are not expected to be encountered with repeated exposures at the PEL, OSHA recommends that employers consider using the **alternative occupational exposure limits** because it believes levels above some of the alternative occupational exposure limits may be hazardous to workers even when the exposure levels are in compliance with the relevant PELs.

Short Term Exposure Limits, Ceiling Limits, and Skin Designations

OSHA also issues Short Term Exposure Limit (STELs) PELs, Ceiling Limits, and PELs that carry a skin designation.

- Short Term Exposure Limit **(STELs)** -- PEL STELs are concentration limits of substances in the air that a worker may be exposed to for 15 minutes without suffering adverse effects. The 15-minute STEL is usually considerably higher than the 8-hour TWA exposure level.
- Ceiling Limits are concentration limits for airborne substances that should never be exceeded.
- A skin designation indicates that the substance can be readily absorbed through the skin, eye or mucous membranes, and substantially contribute to the dose that a worker receives from inhalation of the substance. OSHA standards do not include surface contamination criteria or quantifications for skin absorption.

Theoretically, an occupational substance could have PELs as TWA, STEL, and Ceiling Value, and with a skin designation, but that is rare. Usually, an OSHA-regulated substance will have only a PEL as a time-weighted average.

Immediately Dangerous to Life or Health (IDLH)

The **Immediately Dangerous to Life or Health (IDLH)** occupational exposure guideline was developed jointly by the OSHA and NIOSH Standards Completion Program in 1974. IDLH represents:

- An airborne exposure "likely to cause death or immediate or delayed permanent adverse health effects or prevent escape from such an environment" (NIOSH).

- "An atmosphere that poses an immediate threat to life, would cause irreversible adverse health effects, or would impair an individual's ability to escape from a dangerous atmosphere" (OSHA).

IDLH values can be used in assigning respiratory protection equipment.

Recommended Exposure Limits and Biological Exposure Indices

The NIOSH **Recommended Exposure Limits** (RELs) are also designated as time-weighted average, short-term exposure limits, and ceiling limits. NIOSH also uses immediately dangerous to life or health (IDLH) values.
ACGIH® also has developed **Biological Exposure Indices (BEIs®)** as guidance values for assessing biological monitoring results (concentration of a chemical in biological media such as blood or urine). The OSHA **Policy Statement** on the Uses of TLVs® and BEIs® provides an overview on using these guidelines.

Control Banding (CB)

Control banding **(CB)** is an emerging area internationally for guiding the assessment and management of workplace chemical risks. CB is a technique that determines a control measure such as dilution by air ventilation or engineering controls based on a range or "band" of hazards such as skin irritation or carcinogenic potential and exposures such as an assessment of a small, medium, or large exposure. It is based on the fact that there are a limited number of control approaches, and a history of having many problems solved in the past. CB is used with other health and safety practices such as chemical substitution. It is not a replacement for the use of experts in occupational safety and health and it does not eliminate the need to perform exposure monitoring.

Knowledge Check

1) The Occupational Safety and Health Administration (OSHA) develops legal standards for workplace exposure. These standards are called:

a) Threshold Limit Values (TLVs)
b) Recommended Exposure Limits (RELs)
c) Permissable Exposure Limits (PELs)

Answer

Permissable Exposure Limits (PELs)

OSHA establishes Permissable Exposure Limits, or PELs, which are legal standards for workplace exposure.

2) What exposure standards can be used to assign respiratory protection equipment?

a) Immediately Dangerous to Life or Health (IDLH)
b) Short Term Exposure Limits (STELs)
c) Biological Exposure Indices (BEIs)

Answer

Immediately Dangerous to Life or Health (IDLH)

Immediately Dangerous to Life or Health (IDLH) values can be used in assigning respiratory protection equipment.

This page titled 2.6.6.7.4: Occupational (Workplace) Exposure Standards/Guidelines/Approaches is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **7.4: Occupational (Workplace) Exposure Standards/Guidelines/Approaches** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.8: Basic Physiology

Learning Objectives

After completing this lesson, you will be able to:

- Explain the basic functional and structural organization of the human body.
- Describe homeostasis.
- Explain the role of the four types of tissues in the human body.
- Explain the role of physiological chemicals in normal body functions.

In this section...

Topics include:

2.6.6.8.1: Introduction to Basic Physiology

2.6.6.8.2: Homeostasis

2.6.6.8.3: Organs and Organ Systems

2.6.6.8.4: Tissues

2.6.6.8.5: Cells

2.6.6.8.6: Chemicals

Section 8: Key Points

What We've Covered

This section made the following main points:

- In most cases, toxic substances exert their harmful effects directly on specific cells or biochemicals within the affected organ (specific toxic effects).
- Homeostasis is the ability of the body to maintain relative stability and function despite drastic changes in the external environment or one portion of the body. The primary components of homeostasis include:
  - Stimulus — a change in the environment.
  - Receptor — the site within the body that detects or receives the stimulus.
  - Control center — the operational point at which the signals are received, analyzed, and an appropriate response is determined.
  - Effector — the body site where a response is generated.
  - Feedback mechanisms — methods by which the body regulates the response
- The basic structure and functional organization of the human body is:
  Chemicals → Cells → Tissues → Organs → Organ Systems → Organism
- The human body consists of eleven organ systems.

- Tissues in organs are precisely arranged to work in harmony to perform organ functions.
- There are four types of tissues in the body:
  1. Epithelial tissue protects, absorbs, and secretes substances, and detects sensations.
  2. Connective tissue provides support and holds body tissues together.
  3. Muscle tissue has the ability to contract.
  4. Nerve tissue conducts electrical impulses and conveys information from one area of the body to another.
- The cell is the smallest component of the body that can perform all of the basic life functions.
- Cell components that are most susceptible to cellular damage include the cell membrane, nucleus, ribosomes, peroxisomes, lysosomes, and mitochondria.
- The three categories of physiological chemicals normally functioning in the body are:
  1. Elements — made up of only one atom (examples include hydrogen and oxygen).
  2. Inorganic compounds — simple molecules made up of one or two different elements (examples include water and carbon dioxide).
  3. Organic compounds — contain covalently-bonded carbon and hydrogen and often other elements (examples include DNA, RNA, ATP, and proteins).

This page titled 2.6.6.8: Basic Physiology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

Introduction to Basic Physiology

In order to understand how toxic substances cause harmful changes in organs, tissues, or cells, knowledge of normal physiology and anatomy is needed. This section is an overview of normal physiology, especially as related to the normal body components and how they function. While we show how some xenobiotics can damage the different body components, detailed examples of toxic cellular and biochemical reactions will be covered in later sections.

Complexity of the Body

The body is immensely complex with numerous components, all of which perform precise functions necessary for the body to maintain health and well-being. Malfunction of any component can result in disease or a breakdown of a portion of the body. Toxic substances can damage an organ or organ system so that it cannot function properly, leading to death or sickness of the organism (for example, liver or kidney failure). However, in nearly all cases, the toxic substance actually exerts its harmful effect directly on specific cells or biochemicals within the affected organ. These cell and chemical changes in turn cause the tissue or organ to malfunction.

Specific Toxic Effects

Most toxic substances are usually specific in their toxic damage to particular tissues or organs, referred to as the "target tissues" or "target organs." Toxic effects may affect only a specific type of cell or biochemical reaction. For example:

- The toxic effect of carbon monoxide is due to its binding to a specific molecule (hemoglobin) of a specific cell (red blood cell).
- Organophosphate toxic substances, which inhibit an enzyme (acetylcholinesterase) responsible for modulating neurotransmission at nerve endings.

Systemic Toxic Effects

On the other hand, the effect of some toxic substances may be generalized and potentially damage all cells and thus all tissues and all organs.

- An example is the production of free radicals by whole body radiation. Radiation interacts with cellular water to produce highly reactive free radicals that can damage cellular components. The result can be a range of effects from the death of the cell, to cell malfunction, and to the failure of normal cell division (for example, cancer).
- An example of a multi-organ chemical toxic substance is lead, which damages several types of cells, including kidney cells, nerve cells, and red blood cells.

The body is a remarkable complex living "machine" consisting of trillions of cells and multitudes of biochemical reactions. Each cell has a specific function and cells work together to promote the health and vitality of the organism. The number and types of toxic reactions are likewise very large. While this tutorial cannot possibly present all these types of cellular and biochemical toxic reactions, it is our goal to provide an overview of the primary toxic mechanisms with a few examples that illustrate these mechanisms. It is important to understand that changes at one level in the body can affect homeostasis at several other levels.

This page titled 2.6.6.8.1: Introduction to Basic Physiology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **8.1: Introduction to Basic Physiology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Homeostasis

Homeostasis is the ability of the body to maintain relative stability and function even though drastic changes may take place in the external environment or in one portion of the body. A series of control mechanisms, some functioning at the organ or tissue level and other centrally controlled, maintain homeostasis. The major central homeostatic controls are the nervous and endocrine systems.

Physical and mental stresses, injury, and disease continually challenge us and any of them can interfere with homeostasis. When the body loses its homeostasis, it may plunge out of control, into dysfunction, illness, and even death. Homeostasis at the tissue, organ, organ system, and organism levels reflects the combined and coordinated actions of many cells. Each cell contributes to maintaining homeostasis.

Maintaining Homeostasis

To maintain homeostasis, the body reacts to an abnormal change (induced by a toxic substance, biological organism, or other stress) and makes certain adjustments to counter the change (a defense mechanism). The primary components responsible for the maintenance of homeostasis include:

- **Stimulus** — a change in the environment, such as an irritant, loss of blood, or presence of a foreign chemical.
- **Receptor** — the site within the body that detects or receives the stimulus, senses the change from normal, and sends signals to the control center.
- **Control center** — the operational point at which the signals are received, analyzed, and an appropriate response is determined. This is sometimes referred to as the integration center since it integrates the signals with other information to determine if a response is needed and the nature of a response.
- **Effector** — the body site where a response is generated, which counters the initial stimulus and thus attempts to maintain homeostasis.
- **Feedback mechanisms** — methods by which the body regulates the degree of response that has been elicited. A negative feedback depresses the stimulus to shut off or reduce the effector response, whereas a positive feedback has the effect of increasing the effector response.

Example: Reaction to a Toxin

An example of a homeostatic mechanism can be illustrated by the body's reaction to a toxin that causes anemia and hypoxia (low tissue oxygen) (Figure 1). The production of red blood cells (**erythropoiesis**) is controlled primarily by the hormone, erythropoietin. When the body goes into a state of hypoxia (the stimulus), it prompts the heme protein (the receptor) that signals the kidney to produce erythropoietin (the effector). This, in turn, stimulates the bone marrow to increase red blood cells and hemoglobin, raising the ability of the blood to transport oxygen and thus raises the tissue oxygen levels in the blood and other tissues. This rise in tissue oxygen levels serves to suppress further erythropoietin synthesis (feedback mechanism). In this example, cells and chemicals interact to produce changes that can either disturb homeostasis or restore homeostasis. Toxic substances that damage the kidney can interfere with the production of erythropoietin or toxic substances that damage the bone marrow can prevent the production of red blood cells. This interferes with the homeostatic mechanism described resulting in anemia.

Figure 2.6.6.8.2.1. Homeostatic mechanism to restore levels of red blood cells
(Image Source: NLM)

Knowledge Check

1) The ability of the body to maintain relative stability and function even though drastic changes may take place in the external environment or in one portion of the body is known as:

a) Physiology
b) Homeostasis
c) Toxicity

Answer

Legally enforceable acceptable exposure levels or controls

2) To maintain homeostasis, the body reacts to an abnormal change *(induced by a toxin, biological organism, or other stress)* and makes certain adjustments to counter the change *(a defense mechanism)*. The component of the homeostasis process which detects the change in the environment is known as the:

a) Effector
b) Stimulus
c) Receptor

Answer

Recommended maximum exposure levels which are voluntary and not legally enforceable

This page titled 2.6.6.8.2: Homeostasis is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **8.2: Homeostasis** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.8.3: Organs and Organ Systems

Organ Systems and Organs

Before one can understand how xenobiotics affect these different body components, it's important to understand normal body components and how they function. For this reason, this section provides a basic overview of anatomy and physiology as it relates to toxicity mechanisms.

Basic Body Structure and Organization

We can think of the basic structure and functional organization of the human body as a pyramid or hierarchical arrangement in which the lowest level of organization (the foundation) consists of cells and chemicals. Organs and organ systems represent the highest levels of the body's organization (Figure 1).

Figure 2.6.6.8.3.1. Pyramid represents a hierarchical organization of human body components
(Image Source: NLM)

Simplified definitions of the various levels of organization within the body are:

- **Organ system** — a group of organs that contribute to specific functions within the body. Examples include:
  - Gastrointestinal system
  - Nervous system
- **Organ** — a group of tissues precisely arranged so that they can work together to perform specific functions. Examples include:
  - Liver
  - Brain
- **Tissue** — a group of cells with similar structure and function. There are only four types of tissues:
  1. Epithelial
  2. Connective
  3. Muscle
  4. Nerve
- **Cell** — the smallest living units in the body. Examples include:
  - Hepatocyte
  - Neuron
- **Chemicals** — atoms or molecules that are the building blocks of all matter. Examples include:
  - Oxygen
  - Protein

Organ Systems of the Human Body

The human body consists of eleven organ systems, each of which contains several specific organs. An organ is a unique anatomic structure consisting of groups of tissues that work in concert to perform specific functions. Table 1 includes the structures and functions of these eleven organ systems.

Organ Systems of the Human Body

The human body consists of eleven organ systems, each of which contains several specific organs. An organ is a unique anatomic structure consisting of groups of tissues that work in concert to perform specific functions. Table 1 includes the structures and functions of these eleven organ systems.

Organ System	Functions	Organs
Integumentary	• Barrier to invading organisms and chemicals • Temperature control	• Skin • Hair • Subcutaneous tissue
Skeletal	• Supports and moves body • Protects internal organs • Mineral storage • Blood formation	• Bones • Cartilage • Ligaments • Bone marrow
Muscular	• Locomotion • Heat production	• Muscles • Tendons
Nervous	• Coordinates activities of other organ systems • Responds to sensations	• Brain • Spinal cord • Nerves • Eyes • Ears
Endocrine	• Regulates body functions by chemicals *(hormones)*	• Pituitary gland • Parathyroid gland • Thyroid gland • Adrenal gland • Thymus • Pancreas • Gonads
Cardiovascular	• Transports oxygen and nutrients to tissues • Removes waste products	• Heart • Blood • Blood vessels
Lymphatic	• Returns tissue fluid to blood • Defends against foreign organisms	• Spleen • Lymph nodes • Thymus • Lymphatic vessels
Respiratory	• Oxygen/carbon dioxide exchange	• Lungs • Trachea • Larynx • Nasal cavities • Pharynx
Digestive	• Processes foods • Absorption of nutrients into body	• Stomach • Intestinal tract • Liver • Pancreas • Esophagus • Salivary glands
Urinary	• Elimination of wastes • Regulates pH and volume of blood	• Kidneys • Urinary bladder • Urethra
Reproductive	• Produces germ cells *(eggs and sperm)* • Environment for growth of fetus *(female)*	• Ovaries • Uterus • Mammary glands • Testes

Organ Systems of the Human Body

The human body consists of eleven organ systems, each of which contains several specific organs. An organ is a unique anatomic structure consisting of groups of tissues that work in concert to perform specific functions. Tabl 1 includes the structures and functions of these eleven organ systems.

Organ System	Functions	Organs
Integumentary	• Barrier to invading organisms and chemicals • Temperature control	• Skin • Hair • Subcutaneous tissue
Skeletal	• Supports and moves body • Protects internal organs • Mineral storage • Blood formation	• Bones • Cartilage • Ligaments • Bone marrow
Muscular	• Locomotion • Heat production	• Muscles • Tendons
Nervous	• Coordinates activities of other organ systems • Responds to sensations	• Brain • Spinal cord • Nerves • Eyes • Ears
Endocrine	• Regulates body functions by chemicals *(hormones)*	• Pituitary gland • Parathyroid gland • Thyroid gland • Adrenal gland • Thymus • Pancreas • Gonads
Cardiovascular	• Transports oxygen and nutrients to tissues • Removes waste products	• Heart • Blood • Blood vessels
Lymphatic	• Returns tissue fluid to blood • Defends against foreign organisms	• Spleen • Lymph nodes • Thymus • Lymphatic vessels
Respiratory	• Oxygen/carbon dioxide exchange	• Lungs • Trachea • Larynx • Nasal cavities • Pharynx
Digestive	• Processes foods • Absorption of nutrients into body	• Stomach • Intestinal tract • Liver • Pancreas • Esophagus • Salivary glands
Urinary	• Elimination of wastes • Regulates pH and volume of blood	• Kidneys • Urinary bladder • Urethra
Reproductive	• Produces germ cells *(eggs and sperm)* • Environment for growth of fetus *(female)*	• Ovaries • Uterus • Mammary glands • Testes

Table 2.6.6.8.3.1. Organ systems of the human body

- Prostate gland
- External genitalia

Knowledge Check

1) Groups of cells with similar structure and function are known as:

a) Tissues
b) Organs
c) Organ systems

Answer

Tissues - **This is the correct answer.**

Tissues are groups of cells with similar structure and function. There are only four types of tissues: epithelial tissue, connective tissue, muscle tissue, and nerve tissue.

2) The organ system that transports oxygen and nutrients to tissues and removes waste products is the:

a) Urinary system
b) Integumentary system
c) Cardiovascular system

Answer

Cardiovascular system - **This is the correct answer.**
The cardiovascular system functions to transport oxygen and nutrients to tissues and removes waste products. The primary organs are the heart, blood, and blood vessels.

3) The organ system that regulates body functions by chemicals (hormones) is known as the:

a) Nervous system
b) Reproductive system
c) Endocrine system

Answer

Endocrine system - **This is the correct answer.**
The endocrine system functions to regulate body functions by chemicals (hormones). It contains several organs including the pituitary gland, parathyroid gland, thyroid gland, adrenal gland, thymus, pancreas, and gonads.

This page titled 2.6.6.8.3: Organs and Organ Systems is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **8.3: Organs and Organ Systems** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Tissues

There are four types of tissues dispersed throughout the body, as described below. A type of tissue is not unique for a particular organ and all types of tissue are present in most organs, just as certain types of cells are found in many organs. For example, nerve cells and circulating blood cells are present in virtually all organs.

An "Orchestra" of Tissues

Tissues in organs are precisely arranged so that they can work in harmony in performing organ functions. This is similar to an orchestra that contains various musical instruments, each of which is located in a precise place and contributes exactly at the right time to create harmony. Like musical instruments that are mixed and matched in various types of musical groups, tissues and cells also are present in several different organs and contribute their part to the function of the organ and the maintenance of homeostasis.

Kinds of Tissues in the Body

The four types of tissues are:

1. Epithelial tissue
2. Connective tissue
3. Muscle tissue
4. Nerve tissue

The four types of tissues are similar in that each consists of cells and extracellular materials. However, the types of tissues have different types of cells and differ in the percentage composition of cells and the extracellular materials. Figure 1 illustrates how tissues fit into the hierarchy of body components.

Figure 2.6.6.8.4.1. Hierarchy of body components
(Image Source: NLM)

Epithelial Tissue

Epithelial tissue is specialized to protect, absorb, and secrete substances, as well as detect sensations. It covers every exposed body surface, forms a barrier to the outside world, and controls absorption. Epithelium forms most of the surface of the skin, and the lining of the intestinal, respiratory, and urogenital tracts. Epithelium also lines internal cavities and passageways such as the chest, brain, eye, inner surfaces of blood vessels, the heart, and the inner ear.
Functions of epithelium include:

- Providing physical protection from abrasion, dehydration, and damage by xenobiotics.
- Controlling the permeability of substances in entering or leaving the body.
  - Some epithelia are relatively impermeable; others are readily crossed.
  - Various toxins can damage this epithelial barrier.
- Detecting sensation (sight, smell, taste, equilibrium, and hearing) and conveying this information to the nervous system.
  - For example, touch receptors in the skin respond to pressure by stimulating adjacent sensory nerves.

The epithelium also contains glands and secretes substances such as sweat or digestive enzymes. Others secrete substances into the blood (hormones), such as the pancreas, thyroid, and pituitary gland.
The epithelial cells are classified according to the shape of the cell and the number of cell layers. Three primary cell shapes exist:

squamous (flat), cuboidal, and columnar. There are two types of layering: 1) simple and 2) stratified. Figure 2 illustrates these types of epithelial cells.

Figure 2.6.6.8.4.2. Classification of epithelial tissues
(Source: Adapted from iStock Photos, ©)

Connective Tissue

Connective tissues provide support and hold the body tissues together. They contain more intercellular substances than the other tissues. Connective tissues include blood; bone; cartilage; adipose (fat); and the fibrous and areolar (loose) connective tissues that give support to most organs. The blood and lymph vessels are immersed in the connective tissue media of the body. The blood-vascular system is a component of connective tissue.

In addition to connecting, the connective tissue plays a major role in protecting the body from outside invaders. The hematopoietic tissue is a form of connective tissue responsible for the manufacture of all the blood cells and immunological capability. Phagocytes are connective tissue cells and produce antibodies. If invading organisms or xenobiotics get through the epithelial protective barrier, the connective tissue acts to defend against them.

Figure 2.6.6.8.4.3. Connective tissues
(Image Source: Adapted from Wikimedia Commons, Public Domain)

Muscle Tissue

Muscular tissue is specialized for an ability to contract. Muscle cells are elongated and called muscle fibers. When one end of a muscle cell receives a stimulus, a wave of excitation is conducted through the entire cell so that all parts contract in harmony.
There are three types of muscle cells:

1. **Skeletal muscle** — attached to bone and contracts causing the bones to move.
2. **Cardiac muscle** — contracts to force blood out of the heart and around the body.
3. **Smooth muscle** — can be found in several organs, including the digestive tract, reproductive organs, respiratory tract, and the lining of the bladder. Examples of smooth muscle activity are the:
   - Contraction of the bladder to force urine out.
   - Peristaltic movement to move feces down the digestive system.
   - Contraction of smooth muscle in the trachea and bronchi which decreases the size of the air passageway.

Figure 2.6.6.8.4.4. Muscle tissues
(Image Source: Adapted from iStock Photos, ©)

Nerve Tissue

Nervous tissue is specialized with a capability to conduct electrical impulses and convey information from one area of the body to another. Most of the nervous tissue (98%) is located in the central nervous system, the brain, and spinal cord.

There are two types of nervous tissue: 1) neurons and 2) neuroglia. Neurons (Figure 5) actually transmit the impulses. Neuroglia (Figure 6) provide physical support for the neural tissue, control tissue fluids around the neurons, and help defend the neurons from invading organisms and xenobiotics. Receptor nerve endings of neurons react to various kinds of stimuli (for example, light, sound, touch, and pressure) and can transmit waves of excitation from the farthest point in the body to the central nervous system.

Figure 2.6.6.8.4.5. Human neuron anatomy
(Image Source: Adapted from iStock Photos, ©)

Figure 2.6.6.8.4.6. Types of neuroglia
(Image Source: Adapted from Wikimedia Commons. Blausen.com staff (2014). "Medical gallery of Blausen Medical 2014". *WikiJournal of Medicine* 1 (2). DOI:10.15347/wjm/2014.010. ISSN 2002-4436. View original image.)

Knowledge Check

1) There are only four types of tissues in the body. The type of tissue that is specialized to protect, absorb and secrete substances, detect sensations, covers every exposed body surface, and forms a barrier to the outside world is:

a) Nerve tissue
b) Epithelial tissue
c) Connective tissue
d) Muscle tissue

Answer

Epithelial tissue

Epithelial tissue is specialized to protect, absorb and secrete substances, and detect sensations. It covers every exposed body surface, forms a barrier to the outside world and controls absorption. Epithelium forms most of the surface of the skin, and the lining of the intestinal, respiratory, and urogenital tracts. Epithelium also lines internal cavities and passageways such as the chest, brain, eye, inner surfaces of blood vessels, and heart and inner ear.

This page titled 2.6.6.8.4: Tissues is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **8.4: Tissues** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Cells

Cells are the smallest component of the body that can perform all of the basic life functions. Each cell performs specialized functions and plays a role in the maintenance of homeostasis. While each cell is an independent entity, it is highly affected by damage to neighboring cells. These various cell types combine to form tissues, which are collections of specialized cells that perform a relatively limited number of functions specific to that type of tissue. Several trillion cells make up the human body. These cells are of various types, which can differ greatly in size, appearance, and function.

Primary Cell Components

While there are approximately 200 types of cells, they all have similar features: cell membrane, cytoplasm, organelles, and nucleus. The only exception is that the mature red blood cell does not contain a nucleus. Toxins can injure any of the components of the cell causing cell death or damage and malfunction.
Figure 1 shows the various components of a composite cell.

Figure 2.6.6.8.5.1. Basic cell structure
(Image Source: adapted from National Cancer Institute - SEER Training)

The primary components of a typical cell include the following:

- **Cell membrane** — a phospholipid bilayer which also contains cholesterol and proteins; its functions are to provide support and to control entrance and exit of all materials. We will discuss the structure of the cell membrane and the mechanisms by which chemicals can penetrate or be absorbed into or out of the cell in the **Introduction to Absorption** section later in ToxTutor.
- **Cytoplasm** — a watery solution of minerals, organic molecules, and gases found between the cell membrane and nucleus.
- **Nucleus** — a membrane-bound part of a cell that contains nucleotides, enzymes, and nucleoproteins; the nucleus controls metabolism, protein synthesis, and the storage and processing of genetic information.
- **Cytosol** — the liquid part of the cytoplasm which distributes materials by diffusion throughout the cell.
- **Nucleolus** — a dense region of the nucleus which contains the RNA and DNA; It is the site for rRNA synthesis and assembly of the ribosome components.
- **Endoplasmic reticulum** — an extensive network of membrane-like channels that extends throughout the cytoplasm; it synthesizes secretory products and is responsible for intracellular storage and transport.
- **Ribosomes** — very small structures that consist of RNA and proteins and perform protein synthesis; some ribosomes are fixed (bound to the endoplasmic reticulum) while other ribosomes are free and scattered within the cytoplasm.
- **Mitochondria** — oval organelles bound by a double membrane with inner folds enclosing important metabolic enzymes; they produce nearly all (95%) of the ATP and energy required by the cell.
- **Lysosomes** — vesicles that contain strong digestive enzymes; lysosomes are responsible for the intracellular removal of damaged organelles or pathogens.
- **Peroxisomes** — very small, membrane-bound organelles which contain a large variety of enzymes that perform a diverse set of metabolic functions.
- **Golgi apparatus** — stacks of flattened membranes containing chambers; they synthesize, store, alter, and package secretory products.
- **Centrioles** — there are two centrioles, aligned at right angles, each composed of 9 microtubule triplets; they organize specific fibers of chromosomes during cell division, which move the chromosomes.
- **Cilia** — thread-like projections of the outer layer of the cell membrane, which serve to move substances over the cell surface.

Cell Components Most Susceptible to Xenobiotics

While all components of the cell can be damaged by xenobiotics or body products produced in reaction to the xenobiotics, the components most likely to be involved in cellular damage are the cell membrane, nucleus, ribosomes, peroxisomes, lysosomes, and mitochondria.

Agents that can lead to changes in the permeability of the membrane and the structural integrity of a cell can damage **cell membranes**. The movement of substances through cell membranes is precisely controlled to maintain homeostasis of the cell. Changes in toxin-induced cell membrane permeability may directly cause cell death or make it more susceptible to the entrance of the toxin or to other toxins that follow. The effects in this case may be cell death, altered cell function, or uncontrolled cell division (neoplasia).

Nuclei contain the genetic material of the cell (chromosomes or DNA). Xenobiotics can damage the **nucleus**, which in many cases lead to cell death, by preventing its ability to divide. In other cases, the genetic makeup of the cell may be altered so that the cell loses normal controls that regulate division. That is, it continues to divide and become a neoplasm. How this happens is described in the **Cancer section** of ToxTutor.

Ribosomes use information provided by the nuclear DNA to manufacture proteins. Cells differ in the type of protein they manufacture. For example, the ribosomes of liver cells manufacture blood proteins whereas the ribosomes of fat cells manufacture triglycerides. Ribosomes contain RNA, structurally similar to DNA. Agents capable of damaging DNA may also damage RNA. Thus, toxic damage to ribosomes can interfere with protein synthesis. In the case of damage to liver cell ribosomes, a decrease in blood albumin may result with impairment in the immune system and blood transport.

Lysosomes contain digestive enzymes that normally function in the defense against disease. They can break down bacteria and other materials to produce sugars and amino acids. When xenobiotics damage lysosomes, the enzymes can be released into the cytoplasm where they can rapidly destroy the proteins in the other organelles, a process known as **autolysis**. In some hereditary diseases, the lysosomes of an individual may lack a specific lysosomal enzyme. This can cause a buildup of cellular debris and waste products that is normally disposed of by the lysosomes. In such diseases, known as lysosomal storage diseases, vital cells (such as in heart and brain) may not function normally resulting in the death of the diseased person.

Peroxisomes, which are smaller than lysosomes, also contain enzymes. Peroxisomes normally absorb and neutralize certain toxins such as hydrogen peroxide (H_2O_2) and alcohol. Liver cells contain considerable peroxisomes that remove and neutralize toxins absorbed from the intestinal tract. Some xenobiotics can stimulate certain cells (especially liver) to increase the number and activity of peroxisomes. This, in turn, can stimulate the cell to divide. The xenobiotics that induce the increase in peroxisomes are known as "peroxisome proliferators." Their role in cancer causation are discussed in the **Cancer section** of ToxTutor.

Mitochondria provide the energy for a cell (required for survival), by a process involving ATP synthesis. If a xenobiotic interferes with this process, the death of the cell will rapidly ensue. Many xenobiotics are mitochondrial poisons.

- Examples of poisons that cause cell death by interfering with mitochondria include cyanide, hydrogen sulfide, cocaine, DDT, and carbon tetrachloride.

Knowledge Check

1) There are several different types of cellular organelles. The very small structures *(fixed to the endoplasmic reticulum or free within the cytoplasm)* that consist of RNA and proteins, and function in protein synthesis, are:

a) Nucleus

b) Peroxisomes
c) Lysosomes
d) Ribosomes

Answer

Legally enforceable acceptable exposure levels or controls
Exposure standards are legally enforceable acceptable exposure levels or controls resulting from Congressional or Executive mandate.

2) The organelle that produces nearly all (95%) of the energy required by the cell is the:

a) Nucleolus
b) Golgi apparatus
c) Mitochondria
d) Centioles

Answer

Recommended maximum exposure levels which are voluntary and not legally enforceable
Exposure guidelines are recommended maximum exposure levels that are voluntary and not legally enforceable.

This page titled 2.6.6.8.5: Cells is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **8.5: Cells** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Chemicals

Most toxic effects are initiated by chemical interactions in which a foreign chemical or physical agent interferes with or damages normal chemicals of the body. This interaction results in a body chemical being unable to carry out its function in maintaining homeostasis.
There are many ways this can happen; for example:

- Interference with absorption or disposition of an essential nutrient.
- Interference with nerve transmission (see **Neurotoxicity**).
- Damage to a cell organelle preventing its functioning (see **Cell Damage and Tissue Repair**).

Types of Physiological Chemicals

There are three categories of chemicals normally functioning in the body:

1. **Elements** — substances made up of only one atom; for example:
   - Hydrogen
   - Calcium
   - Singlet oxygen (O)
2. **Inorganic compounds** — simple molecules that usually consist of one or two different elements; for example:
   - Water (H2O)
   - Carbon dioxide (CO2)
   - Bimolecular oxygen (O2)
   - Sodium chloride (NaCl)
3. **Organic compounds** — substances that contain covalently-bonded carbon and hydrogen and often other elements; for example:
   - Sugars
   - Lipids
   - Amino acids
   - Proteins

Elements

Elements are components of all chemical compounds. Of the 92 naturally occurring elements, only 20 are normally found in the body. Seven of these, carbon, oxygen, hydrogen, calcium, nitrogen, phosphorous, and sulfur make up approximately 99% of the human body weight. In most cases, the elements are components of inorganic or organic compounds. However, in a few cases the actual elements may enter into chemical reactions in the body, such as oxygen during cell respiration, sodium in neurotransmission, and arsenic and lead in impaired mitochondrial metabolism.

Inorganic Compounds

Inorganic compounds are important in the body and responsible for many simple functions. The major inorganic compounds are water (H2O), bimolecular oxygen (O2), carbon dioxide (CO2), and some acids, bases, and salts. The body is composed of 60–75% water. Oxygen is required by all cells for cellular metabolism and circulating blood must be well oxygenated for maintenance of

life. Carbon dioxide is a waste product of cells and must be eliminated or a serious change in pH can occur, known as acidosis. A balance in acids, bases, and salts must be maintained to assure homeostasis of blood pH and electrolyte balance.

Organic Compounds

Organic compounds are involved in nearly all biochemical activities related to normal cellular metabolism and function. The mechanisms by which xenobiotics cause cellular and biochemical toxicity are predominantly related to changes to organic compounds. The main feature that differentiates organic compounds from inorganic compounds is that organic compounds always contain carbon. Most organic compounds are also relatively large molecules. There are five major categories of organic compounds involved in normal physiology of the body:

1. Carbohydrates
2. Lipids
3. Proteins
4. Nucleic acids
5. High-energy compounds

Figure 2.6.6.8.6.1. Organic compounds are involved in numerous structures and functions of biochemical processes in the body (Image Source: NLM)

Carbohydrates

Most **carbohydrates** serve as sources of energy for the body. They are converted to glucose, which in turn is used by the cells in cellular respiration. Other carbohydrates become incorporated as structural components of genetic macromolecules.

- For example, deoxyribose is part of DNA, the genetic material of chromosomes, and ribose is part of RNA, which regulates protein synthesis.

Lipids

Lipids are essential substances of all cells and serve as a major energy reserve. They may be stored as fatty acids or as triglycerides. Other types of lipids are the steroids and phospholipids.

- Cholesterol is a lipid that is a component of cell membranes and is used to produce sex hormones such as testosterone and estrogen.
- Phospholipids serve as the main components of the phospholipid bilayer cell membrane.

Proteins

The most diverse and abundant of organic compounds in the body is the group of **proteins**. There are about 100,000 different kinds of proteins, which account for about 20% of the body weight. The building blocks for proteins are the 20 amino acids, which contain carbon, hydrogen, oxygen, nitrogen, and sometimes sulfur. Most protein molecules are large and consist of 50–1000 amino acids bonded together in a very precise structural arrangement. Even the slightest change in the protein molecule alters its function.
Proteins perform a large variety of important functions. Some proteins have a structural function such as the protein pores in cell membranes, keratin in skin and hair, collagen in ligaments and tendons, and myosin in muscles.

- Hemoglobin and albumin are proteins that carry oxygen and nutrients in the circulating blood.
- Antibodies and hormones are proteins.

A particularly important group of proteins are the enzymes.
Enzymes, which are catalysts, are compounds that accelerate chemical reactions, without themselves being permanently changed.

Each enzyme is specific in that it will catalyze only one type of reaction. Enzymes are vulnerable to damage by xenobiotics and many toxic reactions occur by changing the shape of the enzyme ("denaturation") or by inhibiting the enzyme ("inhibition").

Nucleic Acids

Nucleic acids are large organic compounds that store and process information at the molecular level inside virtually all body cells. Three types of nucleic acids are present:

- Deoxyribonucleic acid (DNA)
- Ribonucleic acid (RNA)
- Adenosine triphosphate (ATP)

Nucleic acids are very large molecules composed of smaller units known as nucleotides. A nucleotide consists of a pentose sugar, a phosphate group, and four nitrogenous bases. The sugar in DNA is deoxyribose while the bases are adenine, guanine, cytosine, and thymine. RNA consists of the sugar, ribose, plus the four bases adenine, guanine, cytosine, and uracil. These two types of molecules are known as the molecules of life. For without them, cells could not reproduce and animal reproduction would not occur.

DNA is in the nucleus and makes up the chromosomes of cells. It is the genetic code for hereditary characteristics. RNA is located in the cytoplasm of cells and regulates protein synthesis, using information provided by the DNA. Some toxic agents can damage the DNA causing a mutation, which can lead to the death of the cell, cancer, birth defects, and hereditary changes in offspring. Damage to the RNA causes impaired protein synthesis, responsible for many types of diseases. Figure 2 shows the structure of DNA and RNA. Note that DNA is double-stranded and known as the double helix. RNA is a single strand of nucleotides.

Figure 2.6.6.8.6.2. Structures of DNA and RNA
(Image Source: Adapted from iStock Photos, ©)

High-Energy Compounds

Adenosine triphosphate (ATP) is the most important high-energy compound. It is a specialized nucleotide located in the cytoplasm of cells that serves as a source of cellular energy. ATP contains adenine (amino acid base), ribose (sugar), and three phosphate groups. ATP is created from adenine diphosphate using the energy released during glucose metabolism. One of the phosphates in ATP can later be released along with energy from the broken bond induced by a cellular enzyme.

Knowledge Check

1) A substance in the body that contains covalently-bonded carbon and hydrogen is an:

a) Organic compound
b) Inorganic compound
c) Element

Answer

Organic compound
Organic compounds contain covalently-bonded carbon and hydrogen and often other elements. For example, sugars, lipids, amino acids, and proteins are organic compounds.

2) The nucleic acid located in the nucleus, which makes up the chromosomes of cells, is:

a) ATP
b) RNA
c) DNA

Answer

DNA

DNA is in the nucleus and makes up the chromosomes of cells. It is the genetic code for hereditary characteristics.

This page titled 2.6.6.8.6: Chemicals is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **8.6: Chemicals** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.9: Introduction to Toxicokinetics

Learning Objectives

After completing this lesson, you will be able to:

- Define toxicokinetics.
- Summarize the four inter-related processes of toxicokinetics.
- Identify examples of transporter proteins and their role in toxicokinetics.

In this section...

Topics include:

2.6.6.9.1: What is Toxicokinetics

What We've Covered

This section made the following main points:

- Toxicokinetics is essentially the study of how a substance enters the body and what happens to it inside the body.
  - The term "disposition" is often used in place of toxicokinetics to describe how the body disposes of a xenobiotic over time.
- The four inter-related processes of toxicokinetics are:
  1. Absorption — the substance enters the body.
  2. Distribution — the substance moves from the site of entry to other areas of the body.
  3. Biotransformation — the substance is transformed into new chemicals (metabolites).
  4. Excretion — the substance or its metabolites leave the body.
- The disposition of a toxicant and its biological reactivity are the factors that determine the severity of toxicity when a xenobiotic enters the body.

This page titled 2.6.6.9: Introduction to Toxicokinetics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

What is Toxicokinetics?

Toxicokinetics Defined

Toxicokinetics is essentially the study of "how a substance gets into the body and what happens to it in the body." Before this term was used, the study of the kinetics (movement) of chemicals was originally conducted with pharmaceuticals and the term pharmacokinetics became commonly used. Similarly, toxicology studies were initially conducted with drugs. Toxicokinetics deals with what the body does with a drug when given a relatively high dose relative to the therapeutic dose. Read more about differences between pharmacokinetics and toxicokinetics.

Processes

Four processes are involved in toxicokinetics:

1. **Absorption** — the substance enters the body.
2. **Distribution** — the substance moves from the site of entry to other areas of the body.
3. **Biotransformation** — the body changes (transforms) the substance into new chemicals (metabolites).
4. **Excretion** — the substance or its metabolites leave the body.

The science of toxicology has evolved to include environmental and occupational chemicals as well as drugs. Toxicokinetics is thus the appropriate term for the study of the kinetics of all substances at toxic dose/exposure levels.

Frequently the terms toxicokinetics, pharmacokinetics, or disposition have the same meaning. Disposition is often used in place of toxicokinetics to describe the movement of chemicals through the body over the course of time, that is, how the body disposes of a xenobiotic.

Figure 2.6.6.9.1.1. Processes of toxicokinetics
(Image Source: Adapted from iStock Photos, ©)

Factors Determining the Severity of Toxicity

The disposition of a toxicant and its biological reactivity are the factors that determine the severity of toxicity that results when a xenobiotic enters the body. The most important aspects of disposition include:

- **Duration and concentration** of a substance at the portal of entry.
- **Rate and amount** of the substance that can be absorbed.
- **Distribution** in the body and **concentration** of the substance at specific body sites.
- **Efficiency** of biotransformation and nature of the metabolites.
- **Ability** of the substance or its metabolites **to pass through cell membranes** and come into contact with specific cell components (for example, DNA).
- **Amount and duration of storage** of the substance (or its metabolites) in body tissues.
- **Rate and sites of excretion** of the substance.
- **Age and health status** of the person exposed.

Here are some examples of how toxicokinetics of a substance can influence its toxicity:

- **Absorption** — A highly toxic substance that is poorly absorbed may be no more hazardous than a substance of low toxicity that is highly absorbed.

- **Biotransformation** — Two substances with equal toxicity and absorption may differ in how hazardous they are depending on the nature of their biotransformation. A substance that is biotransformed into a more toxic metabolite (bioactivated) is a greater hazard than a substance that is biotransformed into a less toxic metabolite (detoxified).

Inter-Related Processes of Absorption, Distribution, Biotransformation, and Elimination

Absorption, distribution, biotransformation, and elimination are inter-related processes as illustrated in Figure 2 below. After the substance is absorbed, it is distributed through the blood, lymph circulation, and extracellular fluids into organs or other storage sites and may be metabolized. Then, the substance or its metabolites are eliminated through the body's waste products.

Figure (\PageIndex{2}\). Absorption, Distribution, Metabolism, and Elimination
(Image Source: NLM)

What are Transporters?

Transporters, also called transporter proteins, play an important role in the processes of absorption, distribution, metabolism, and elimination (ADME). They are important to pharmacological, toxicological, clinical, and physiological applications. For example:

- In the **liver** — transmembrane transporters, together with drug metabolizing enzymes, are important in drug metabolism and drug clearance by the liver. Xenobiotics, endogenous metabolites, bile salts, and cytokines affect the levels (or "expression") of these transporters in the liver. Adverse reactions in the liver to a xenobiotic such as a drug could be caused by genetic or disease-induced variations of transporter expression or drug-drug interactions at the level of these transporters.
- In the **kidneys** — renal proximal tubules are targets for toxicity partly because of the expression of transporters that mediate the secretion and reabsorption of xenobiotics. Changes in transporter expression and/or function could enhance the accumulation of toxicants and make the kidneys more susceptible to injury, for example, when xenobiotic uptake by carrier proteins is increased or the efflux of toxicants and their metabolites is reduced. The list of nephrotoxic chemicals is a long one and includes:
  - Environmental contaminants such as some hydrocarbon solvents, some heavy metals, and the fungal toxin ochratoxin.
  - Some antibiotics.
  - Some antiviral drugs.
  - Some chemotherapeutic drugs.

The competition of xenobiotics for transporter-related excretion and genetic polymorphisms affecting transporter function affect the likelihood of nephrotoxicity.
Because of concerns that such changes to transporter expression and function can adversely affect clinical outcomes and physiological regulation, increased drug transporter activity is important to study and understand. There is clinical and laboratory research including *in vitro*, *ex vivo*, and *in vivo* studies that shows how powerful drug-drug interactions can be.
For example, drugs might compete with each other for binding to a transporter, which can lead to changes in serum and tissue drug levels and possible side effects.

- This is one possible explanation for the rare occurrence of potentially severe toxicity when the drug methotrexate and nonsteroidal anti-inflammatory drugs are given at the same time.
- The drug probenecid, which competitively inhibits some transporters, has been used to increase the half-life of antibiotics such as penicillin and antiviral drugs and improve their therapeutic value.

Pharmacokinetics and Toxicokinetics: Now and in the Future

Current research priorities suggest that we can anticipate important strides in the following areas of pharmacokinetics and toxicokinetics:

- An increased understanding of human variability of pharmacokinetics and pharmacodynamics in the population.
- Further exploration of mode of action hypotheses (**MoA**).
    - Is a MoA the same as a MOA? No. A **mode of action (MoA)** describes a functional or anatomical change, at the cellular level, resulting from exposure to a substance. A **mechanism of action (MOA)** describes changes at the molecular level.
- Further application of biological modeling in the risk assessment of individual chemicals and chemical mixtures.
- Further identification and discussion of uncertainties in the modeling process.
- Further use of "Reverse Toxicokinetics," also called **"IVIVE" (*In vitro* to *in vivo* extrapolation)**. IVIVE *in vitro* data to estimate exposures that could be associated with adverse effects *in vivo*.

This page titled 2.6.6.9.1: What is Toxicokinetics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **9.1: What is Toxicokinetics** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.10: Absorption

Learning Objectives

After completing this lesson, you will be able to:

- Explain absorption and its role in toxicokinetics.
- Describe the primary routes of exposure.
- Explain the role of cell membranes in absorption.
- Identify ways in which xenobiotics pass across cell membranes.

In this section...

Topics include:

2.6.6.10.1: Introduction to Absorption

2.6.6.10.2: Gastrointestinal Tract

2.6.6.10.3: Respiratory Tract

2.6.6.10.4: Dermal Route

2.6.6.10.5: Other Routes of Exposure

Section 10: Key Points

What We've Covered

This section made the following main points:

- Absorption is the process by which toxicants gain entrance into the body.
- Ingested and inhaled materials are considered outside the body until they cross the cellular barriers of the gastrointestinal tract or respiratory system.
- The likelihood of absorption depends on the:
  - Route of exposure.
  - Concentration of the substance at the site of contact.
  - Chemical and physical properties of the substance.
- Exposure routes include:
  - Primary routes:
    - Gastrointestinal (GI) tract
      - Mouth and esophagus — poorly absorbed under normal conditions due to short exposure time (nicotine and nitroglycerin are notable exceptions).
      - Stomach — significant site for absorption of weak organic acids, but weak bases are poorly absorbed.
      - Intestine — greatest absorption of both weak bases and weak acids, particularly in the small intestine.

        - Colon and rectum — very little absorption, unless administered via suppository.
    - Respiratory tract
        - Mucociliary escalator — movements of the cilia push mucus and anything contained within up and out into the throat to be swallowed or removed through the mouth.
        - Pulmonary region — most important site for absorption with about 50 times the surface area of the skin and very thin membranes.
    - Skin
        - Epidermis and stratum corneum — the only layer important in regulating the penetration of a skin contaminant.
        - Toxicants move across the stratum corneum by passive diffusion.
        - If a toxicant penetrates through the stratum corneum, it enters lower layers of the epidermis, dermis, and subcutaneous tissue, which are far less resistant to further diffusion.
  - Other exposure routes:
    - Injections
    - Implants
    - Conjunctival instillations (eye drops)
    - Suppositories
- Cell membranes surround all body cells and are made up of a phospholipid bilayer in which each molecule contains a:
  - Polar (hydrophilic, or attracted to water) phosphate head
  - Lipophilic (attracted to lipid-soluble substances) lipid tail
- Xenobiotics must pass across cell membranes to enter, move within, and leave the body. This movement can be either:
  - Passive transfer (most common) — simple diffusion or osmotic filtration with no cellular energy or assistance required.
  - Facilitated transport — similar to passive transport, but a carrier-mediated transport mechanism and thus faster and capable of moving larger molecules.
  - Active transport — movement against the concentration gradient (from lower to higher concentrations), requiring cellular energy from ATP.
  - Endocytosis — the cell surrounds the substance with a section of its cell wall, separating from the membrane and moving into the interior of the cell.

This page titled 2.6.6.10: Absorption is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

Introduction to Absorption

Toxicants gain entrance into the body by absorption. The body considers ingested and inhaled materials as being outside of it until those materials cross the cellular barriers of the gastrointestinal tract or respiratory system. A substance must be absorbed to exert an effect on internal organs, although local toxicity, such as irritation, may occur.

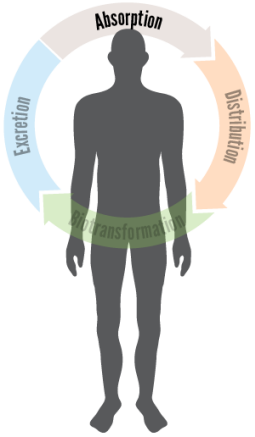


Figure 2.6.6.10.1.1. Processes of toxicokinetics
(Image Source: Adapted from iStock Photos, ©)

Absorption Variability

Absorption varies greatly by specific chemicals and the route of exposure.

- For skin, oral or respiratory exposure, the absorbed dose is only a fraction of the exposure dose (external dose).
- For substances injected or implanted directly into the body, exposure dose is the *same* as the absorbed or internal dose.

Several factors affect the likelihood that a xenobiotic will be absorbed. The most important factors are the:

- Route of exposure.
- Concentration of the substance at the site of contact.
- Chemical and physical properties of the substance.

The route of exposure influences how the concentration and properties of the substance vary. In some cases, a high percentage of a substance may not be absorbed from one route whereas a low amount may be absorbed via another route.

- For example, very little DDT powder will penetrate the skin whereas a high percentage will be absorbed when it is swallowed.

Due to such route-specific differences in absorption, xenobiotics are often ranked for hazard in accordance with the route of exposure. A substance may be categorized as relatively non-toxic by one route and highly toxic via another route.

Routes of Exposure

The **primary routes of exposure** by which xenobiotics can gain entry into the body are:

- **Gastrointestinal (GI) tract** — important for environmental exposure to contaminants from food and water; the main route for many pharmaceuticals.
- **Respiratory tract** — important for environmental and occupational exposure to air contaminants; some pharmaceuticals (such as nasal or oral aerosol inhalers) use this route.
- **Skin** — important environmental and occupational exposure route; many consumer and pharmaceutical products are applied directly to the skin.

Other routes of exposure – used primarily for specific medical purposes:

- **Injections** — primarily used for pharmaceuticals.
- **Implants** — pharmaceuticals may be implanted to permit slow, time-release (for example, hormones). Many medical devices are implanted for which minimal absorption is desired (such as artificial lens or tendons). Some materials enter the body via skin penetration as the result of accidents or weapons.
- **Conjunctival instillations (eye drops)** — primarily for treating ocular conditions; however, in some cases, considerable absorption can occur and cause systemic toxicity.
- **Suppositories** — used for medicines that may not be adequately absorbed after oral administration or that are intended for local therapy; usual locations for suppositories are the rectum and vagina.

Cell Membranes

Cell membranes (often referred to as plasma membranes) surround all body cells and are similar in structure. They consist of two layers of phospholipid molecules arranged like a sandwich, referred to as a "phospholipid bilayer." Each phospholipid molecule consists of a phosphate head and a lipid tail. The phosphate head is polar, meaning it is hydrophilic *(attracted to water)*. In contrast, the lipid tail is lipophilic *(attracted to lipid-soluble substances)*.
The two phospholipid layers are oriented on opposing sides of the membrane so that they are approximate mirror images of each other. The polar heads face outward and the lipid tails face inward in the membrane sandwich (Figure 2.6.6.10.1.2).

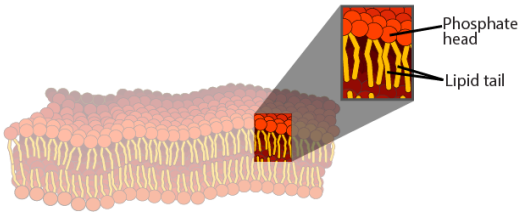


Figure 2.6.6.10.1.2. Each phospholipid molecule consists of a phosphate head and lipid tail
(Image Source: Adapted from Wikimedia Commons, obtained under Public Domain, original image)

The cell membrane is tightly packed with these phospholipid molecules interspersed with various proteins and cholesterol molecules. Some proteins span across the entire membrane which can create openings for aqueous channels or pores.

A typical cell membrane structure is illustrated in Figure 2.6.6.10.1.3.

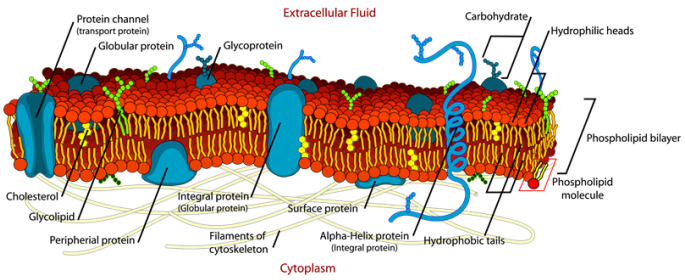


Figure 2.6.6.10.1.3. Typical cell membrane structure (Image Source: Adapted from Wikimedia Commons, obtained under Public Domain, original image)

Role of Cell Membranes in Absorption

For a xenobiotic to enter the body (as well as move within, and leave the body) it must pass across cell membranes (cell walls). Cell membranes are formidable barriers and a major body defense that prevents foreign invaders or substances from gaining entry into body tissues. Normally, cells in solid tissues (such as skin, or mucous membranes of the lung or intestine) are so tightly compacted that substances cannot pass between them. This requires that the xenobiotic have the ability to penetrate cell membranes. It must cross several membranes to go from one area of the body to another.

For a substance to move through one cell requires that it first move across the cell membrane into the cell, pass across the cell, and then cross the cell membrane again to leave the cell. This is true whether the cells are in the skin, the lining of a blood vessel, or an internal organ such as the liver. In many cases, in order for a substance to reach the site where it exerts toxic effects, it must pass through several membrane barriers.

Animation 1 depicts how a chemical from a theoretical consumer product called a "Shower Gel" might get to the surface of the skin during showering and then pass through several membranes before coming in contact with the inside of a liver cell

http://www.toxmsdt.com/uploads/b/886...a_cell_446.mp4

Animation 2.6.6.10.1.1. From a Gel to a Cell: Following the journey of a chemical from a theoretical shower gel product through several membranes and ultimately into a cell
(Image Source: iStock Photos, ©)

Movement of Toxicants Across Cell Membranes

Some toxicants move across a membrane barrier with relative ease while others find it difficult or impossible. Those that can cross the membrane use one of two general methods: 1) **passive transfer** or 2) **facilitated transport**.
Passive transfer consists of **simple diffusion** (or osmotic filtration) and is "passive" because no cellular energy or assistance is required.
Some toxicants cannot simply diffuse across the membrane but require assistance by specialized transport mechanisms. The primary types of specialized transport mechanisms are:

- Facilitated diffusion
- Active transport
- Endocytosis (phagocytosis and pinocytosis)

Passive Transfer

Passive transfer is the most common way that xenobiotics cross cell membranes. Two factors determine the rate of passive transfer:

1. The difference in concentrations of the substance on opposite sides of the membrane (this occurs when a substance moves from a region of high concentration to one having a lower concentration. Diffusion will continue until the concentration is equal on both sides of the membrane).
2. The ability of the substance to move either through the small pores in the membrane or the lipophilic interior of the membrane.

Properties affecting a chemical substance's ability for passive transfer are:

- Lipid solubility
- Molecular size

- The degree of ionization

Substances with high lipid solubility readily diffuse through the phospholipid membrane. Small water-soluble molecules can pass across a membrane through the aqueous pores, along with normal intracellular water flow.
Large water-soluble molecules usually cannot make it through the small pores, although some may diffuse through the lipid portion of the membrane, but at a slow rate.

- Most aqueous pores are about 4 Angstrom (Å) in size and allow chemicals of molecular weight 100-200 to pass through. Exceptions are membranes of capillaries and kidney glomeruli which have relatively large pores (about 40 Angstrom [Å]) that allow molecules up to a molecular weight of about 50,000 (molecules slightly smaller than albumin which has a molecular weight of 60,000) to pass through.

In general, highly ionized chemicals have low lipid solubility and pass with difficulty through the lipid membrane.
Figure 4 demonstrates the passive diffusion and filtration of xenobiotics through a typical cell membrane.

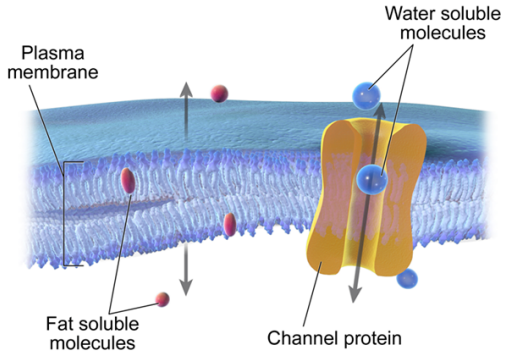


Figure 2.6.6.10.1.4. Cellular diffusion
(Image Source: Blausen.com staff (2014). "Medical gallery of Blausen Medical 2014". WikiJournal of Medicine 1 (2). DOI:10. 15347/wjm/2014.010. ISSN 2002-4436. Obtained under Creative Commons license.)

Facilitated Diffusion

Facilitated diffusion is similar to simple diffusion in that it does not require energy and follows a concentration gradient. The difference is that it is a carrier-mediated transport mechanism (Figure 2.6.6.10.1.5)—that is, special transport proteins, which are embedded within the cell membrane, facilitate movement of molecules across the membrane. The results are similar to passive transport but faster and capable of moving larger molecules that have difficulty diffusing through the membrane without a carrier.

- Examples are the transport of sugar and amino acids into RBCs and the CNS.

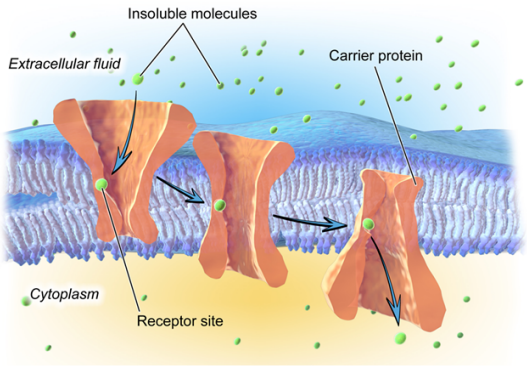


Figure 2.6.6.10.1.5. Facilitated Diffusion
(Image Source: Blausen.com staff (2014). "Medical gallery of Blausen Medical 2014". WikiJournal of Medicine 1 (2). DOI:10.15347/wjm/2014.010. ISSN 2002-4436. Obtained under Creative Commons license.)

Active Transport

Some substances are unable to move with diffusion, unable to dissolve in the lipid layer, and are too large to pass through the aqueous channels. For some of these substances, **active transport** processes exist in which movement through the membrane may be *against* the concentration gradient, that is, from low to higher concentrations. Cellular energy from adenosine triphosphate (ATP) is required in order to accomplish this. The transported substance can move from one side of the membrane to the other side by this energy process. Active transport is important in the transport of xenobiotics into the liver, kidney, and central nervous system and for maintenance of electrolyte and nutrient balance.

Figure 2.6.6.10.1.6 shows sodium and potassium moving against concentration gradient with the help of the ATP sodium-potassium exchange pump.

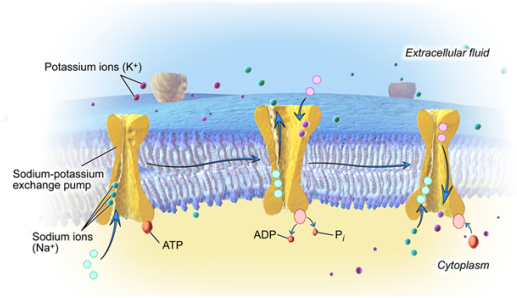


Figure 2.6.6.10.1.6. Sodium-Potassium Exchange Pump
(Image Source: Blausen.com staff (2014). "Medical gallery of Blausen Medical 2014". WikiJournal of Medicine 1 (2). DOI:10.15347/wjm/2014.010. ISSN 2002-4436. Obtained under Creative Commons license.)

Endocytosis (Phagocytosis and Pinocytosis)

Many large molecules and particles cannot enter cells via passive or active mechanisms. However, some may still enter by a process known as **endocytosis**.
In endocytosis, the cell surrounds the substance with a section of its cell wall. This engulfed substance and section of membrane then separates from the membrane and moves into the interior of the cell. The two main forms of endocytosis are 1) phagocytosis and 2) pinocytosis.
In **phagocytosis** (cell eating), large particles suspended in the extracellular fluid are engulfed and either transported into cells or are destroyed within the cell. This is a very important process for lung phagocytes and certain liver and spleen cells.
Pinocytosis (cell drinking) is a similar process but involves the engulfing of liquids or very small particles that are in suspension in the extracellular fluid.

Figure 2.6.6.10.1.7 demonstrates the types of membrane transport by endocytosis.

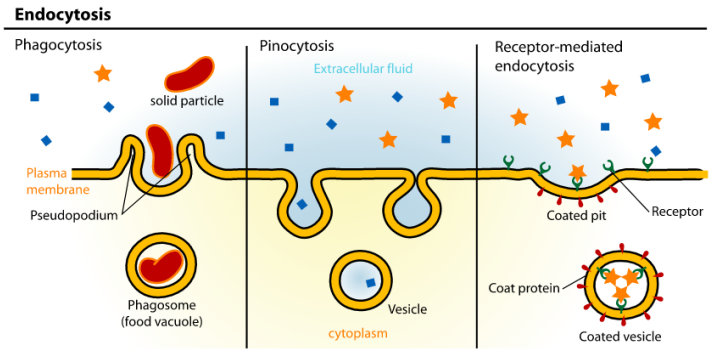


Figure 2.6.6.10.1.7. Types of Endocytosis
(Image Source: Wikimedia Commons, obtained by Public Domain License, original image)

Knowledge Check

1) The process whereby a substance moves from outside the body into the body is known as:
a) Distribution
b) Biotransformation
c) Absorption

Answer

Absorption - **This is the correct answer.**
Absorption is the first and crucial step in the toxicokinetics of a xenobiotic. Without absorption, a toxic substance does not represent a human health hazard.

2) For a xenobiotic to move from outside the body to a site of toxic action requires that it:
a) Possess hydrophilic (water-soluble) properties
b) Possess hydrophobic (lipophilic) properties
c) Pass through several cell membranes

Answer

Pass through several cell membranes - **This is the correct answer.**
In order for a xenobiotic to move from outside the body to an internal site of toxic action (target cells), a xenobiotic must pass through several membrane barriers. The first membranes are those at the portal of entry, for example, lung or intestinal tract.

3) The basic structure of the cell membrane consists of:
a) A thick protein layer containing phospholipid channels
b) A bilayer of phospholipids with scattered proteins within the layers
c) Cholesterol outer layer with a phospholipid inner layer

Answer

A bilayer of phospholipids with scattered proteins within the layers - **This is the correct answer.**
The typical cell membrane consists of two layers of phospholipids with polar head groups consisting of phospholipid molecules and the lipid inner portion consisting primarily of cholesterol molecules. The phospholipid layers are oriented on opposing sides of the membrane so that they are approximate mirror images of each other. Various proteins are scattered throughout the lipid bilayers of the membrane.

4) The membrane transport process by which large hydrophobic molecules cross membranes via the lipid portion of the membrane, follow the concentration gradient, and do not require energy or carrier molecules is known as:

a) Simple diffusion
b) Active transport
c) Facilitated diffusion

Answer

Simple diffusion - **This is the correct answer.**
Large hydrophobic molecules must diffuse through the lipid portion of the membrane, with the rate of transport correlating with its lipid solubility. In general, highly ionized chemicals have low lipid solubility and do not readily pass through the lipid membrane.

5) Endocytosis is a form of specialized membrane transport in which the cell surrounds the substance with a section of its cell membrane. The specific endocytosis process by which liquids or very small particles are engulfed and transported across the membrane is known as:

a) Phagocytosis
b) Pinocytosis
c) Exocytosis

Answer

Pinocytosis - **This is the correct answer.**
Pinocytosis (cell drinking) involves the engulfing of liquids or very small particles that are in suspension within the extracellular fluid.

This page titled 2.6.6.10.1: Introduction to Absorption is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **10.1: Introduction to Absorption** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Gastrointestinal Tract

The gastrointestinal (GI) tract can be viewed as a tube going through the body (Figure 2.6.6.10.2.1). Its contents are considered exterior to the body until absorbed. Salivary glands, liver, and the pancreas are considered accessory glands of the GI tract as they have ducts entering the GI tract and secrete enzymes and other substances. For foreign substances to enter the body, they must pass through the gastrointestinal mucosa, crossing several membranes before entering the bloodstream.

Substances must be absorbed from the gastrointestinal tract in order to exert a toxic effect throughout the whole body, although local gastrointestinal damage may occur from direct exposures to toxicants. Absorption can occur at any place along the entire gastrointestinal tract. However, the degree of absorption depends on the site.

Three main factors affect absorption within the various sites of the gastrointestinal tract:

1. Type of cells at the specific site.
2. Period of time that the substance remains at the site.
3. pH of stomach or intestinal contents at the site.

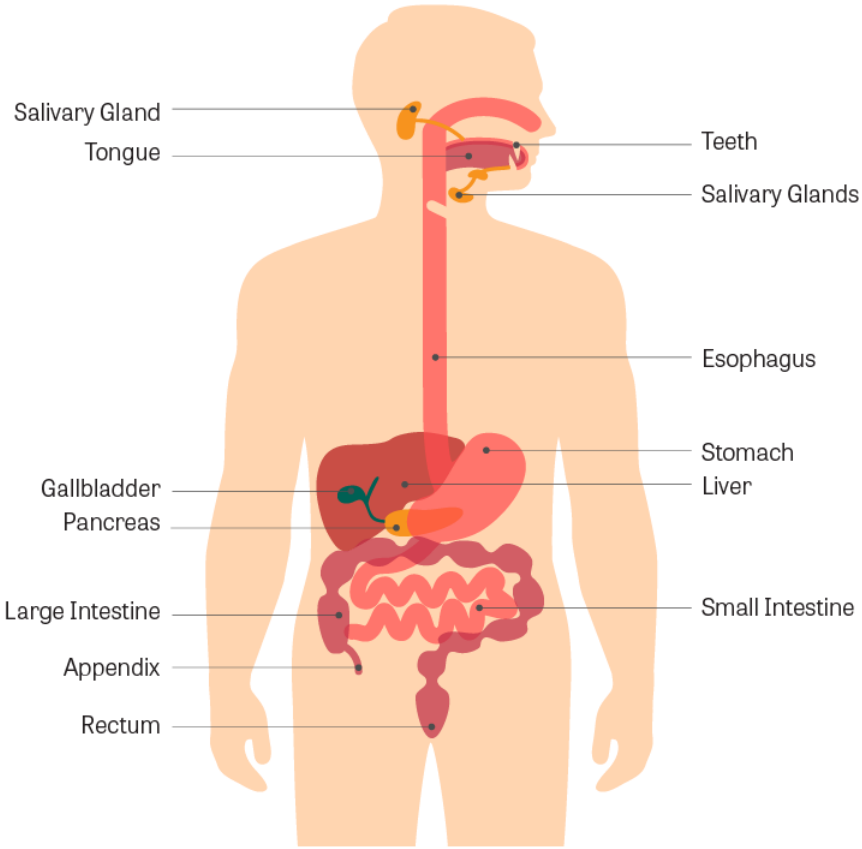


Figure 2.6.6.10.2.1. Anatomy of gastrointestinal tract
(Image Source: adapted from iStock Photos, ©)

Mouth and Esophagus

Under normal conditions, xenobiotics are poorly absorbed within the **mouth** and **esophagus**, due mainly to the very short time that a substance resides within these portions of the gastrointestinal tract. There are some notable exceptions. For example:

- Nicotine readily penetrates the mouth mucosa.
- Nitroglycerin is placed under the tongue (sublingual) for immediate absorption and treatment of heart conditions.

The sublingual mucosa under the tongue and in some areas of the mouth is thin and highly vascularized and allows some substances to be rapidly absorbed.

Stomach

The **stomach**, with its high acidity (pH 1-3), is a significant site for the absorption of weak organic acids, which exist in a diffusible, nonionized and lipid-soluble form. In contrast, weak bases will be highly ionized and therefore poorly absorbed. The acidic stomach may chemically break down some substances. For this reason, those substances must be administered in gelatin capsules or coated tablets, which can pass through the stomach into the intestine before they dissolve and release their contents.

Another determinant that affects the amount of a substance that will be absorbed in the stomach is the presence of food in the stomach. Food ingested at the same time as the xenobiotic may result in a considerable difference in absorption of the xenobiotic.

Intestine

The greatest absorption of chemicals, as with nutrients, takes place in the **intestine**, particularly in the small intestine. The intestine has a large surface area consisting of outward projections of the thin (one-cell thick) mucosa into the lumen of the intestine (the villi) (Figure 2.6.6.10.2.2). This large surface area facilitates diffusion of substances across the cell membranes of the intestinal mucosa.

Since the pH is near neutral (pH 5-8), both weak bases and weak acids are nonionized and are usually readily absorbed by passive diffusion. Lipid soluble, small molecules effectively enter the body from the intestine by passive diffusion.

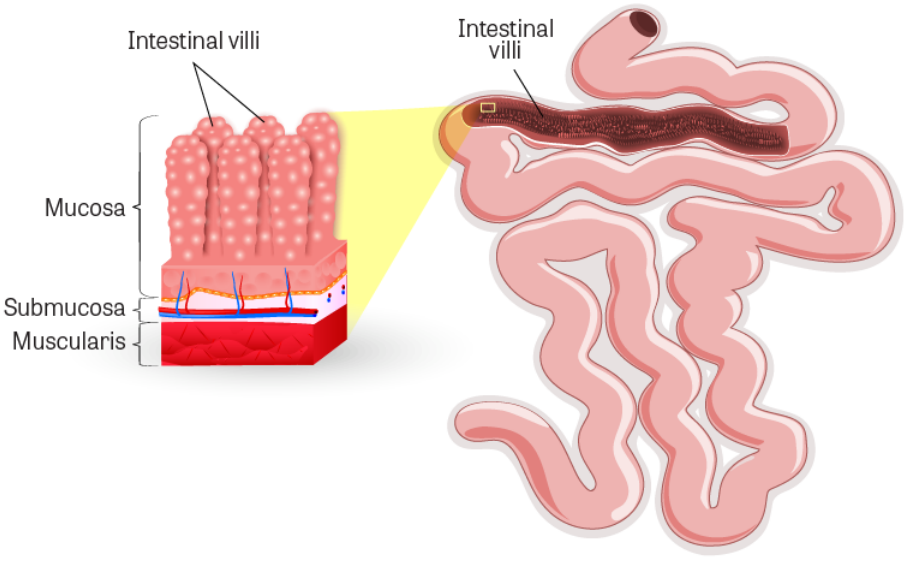


Figure 2.6.6.10.2.2. Anatomy of structures in small intestine used for absorption
(Image Source: adapted from iStock Photos, ©)

In addition to passive diffusion, facilitated and active transport mechanisms move certain substances across the intestinal cells into the body, including essential nutrients such as glucose, amino acids, and calcium. These mechanisms also transport strong acids, strong bases, large molecules, and metals, including some important toxins.

- For example, lead, thallium, and paraquat (herbicide) are toxins that active transport systems move across the intestinal wall.

The slow movement of ingested substances through the intestinal tract can influence their absorption. This slow passage increases the length of time that a compound is available for absorption at the intestinal membrane barrier.

Intestinal microflora and gastrointestinal enzymes can affect the toxicity of ingested substances. Some ingested substances may be only poorly absorbed but they may be biotransformed within the gastrointestinal tract. In some cases, their biotransformed products may be absorbed and be more toxic than the ingested substance.

- An important example is the formation of carcinogenic nitrosamines from non-carcinogenic amines by intestinal flora.

Colon and Rectum

Very little absorption takes place in the **colon** and **rectum**. As a rule, if a xenobiotic has not been absorbed after passing through the stomach or small intestine, very little further absorption will occur. However, there are some exceptions, as some medicines may be administered as rectal suppositories with significant absorption.

- An example is Anusol (hydrocortisone preparation) used for the treatment of local inflammation which is partially absorbed (about 25%).

Knowledge Check

1) The most important factor that determines whether a substance will be absorbed within the stomach is the:

a) Physical form as a solid or liquid
b) Molecular size
c) pH

Answer

pH - **This is the correct answer.**
The most important factor that determines absorption within the stomach is pH. Weak organic acids, which exist in a diffusible, nonionized and lipid-soluble form are readily absorbed in the high acidity of the stomach (pH 1-3). In contrast, weak bases will be highly ionized and therefore poorly absorbed.

2) The primary routes for absorption of environmental agents are:

a) Gastrointestinal tract, respiratory tract, and skin
b) Conjunctival exposures and skin wounds

Answer

Gastrointestinal tract, respiratory tract, and skin - **This is the correct answer.**
Environmental agents may be found in contaminated food, water, or air. As such, they may be ingested, inhaled, or present on the skin.

3) The site of the gastrointestinal tract where most absorption takes place is:

a) Stomach

b) Small intestine
c) Colon and rectum

Answer

Small intestine - **This is the correct answer.**
By far, the greatest absorption takes place in the intestine. This is due to the neutral pH and the large, thin, surface area that allows easy penetrable by passive diffusion. Weak bases, weak acids, lipid soluble substances and small molecules effectively enter the body from the intestine. In addition, special carrier-mediated and active transport systems exist.

This page titled 2.6.6.10.2: Gastrointestinal Tract is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **10.2: Gastrointestinal Tract** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Respiratory Tract

Many environmental and occupational agents as well as some pharmaceuticals enter the respiratory tract through inhalation. Absorption can occur at any place within the upper respiratory tract. However, the amount of a particular xenobiotic that can be absorbed at a specific location depends highly on its physical form and solubility.
There are three basic regions to the respiratory tract:

1. Nasopharyngeal region
2. Tracheobronchial region
3. Pulmonary region

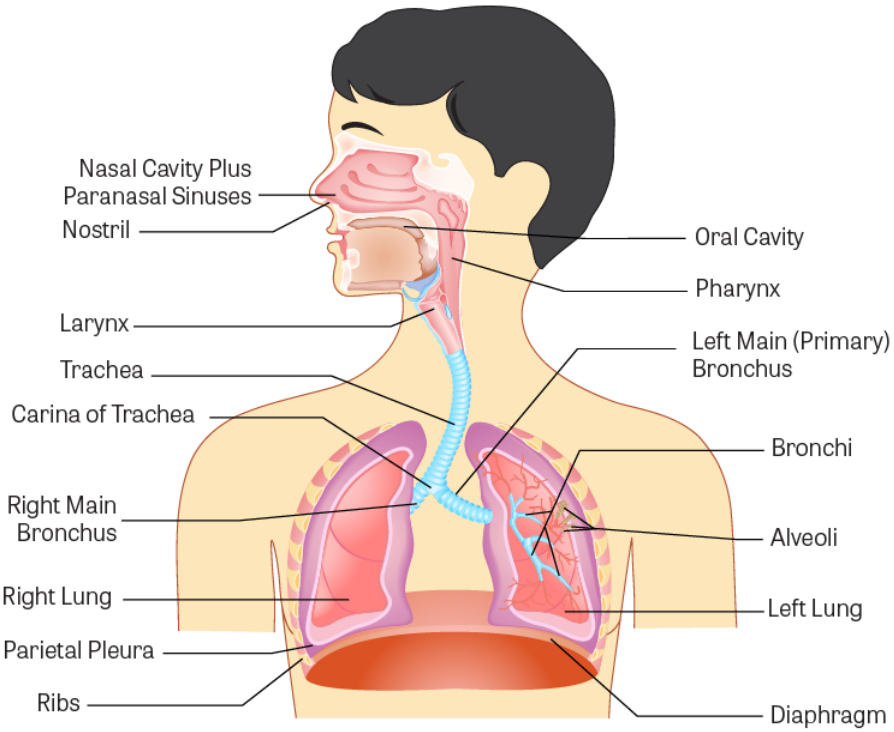
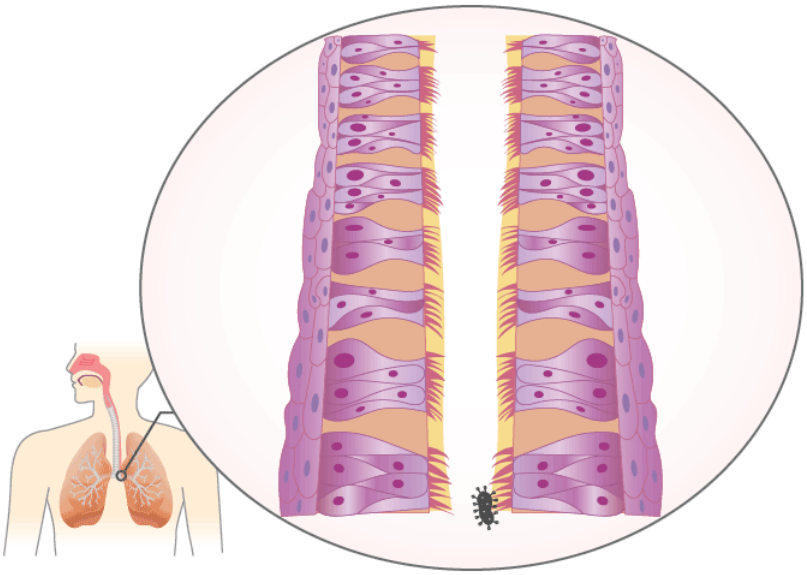


Figure 2.6.6.10.3.1. Anatomy of the respiratory tract
(Image Source: adapted from iStock Photos, ©)

Mucociliary Escalator

The mucociliary escalator covers most of the bronchi, bronchioles, and nose. It contains mucus-producing goblet cells and ciliated epithelium. The movements of the cilia push it and anything in it such as inhaled particles or microorganisms up and out into the throat, which can either get swallowed or removed through the mouth.



Animation 2.6.6.10.3.1. The mucociliary escalator provides a barrier against infection

Pulmonary Region

By far, the most important site for absorption is the pulmonary region consisting of the very small airways called bronchioles and the alveolar sacs of the lung.

The alveolar region has a very large surface area, about 50 times that of the skin. In addition, the alveoli consist of only a single layer of cells with very thin membranes that separate the inhaled air from the blood stream. Oxygen, carbon dioxide, and other gases readily pass through this membrane. Gases and particles, which are water-soluble (and thus blood-soluble), are absorbed more efficiently from the lung alveoli compared to their absorption via the gastrointestinal tract or through the skin. Water-soluble gases and liquid aerosols can pass through the alveolar cell membrane by simple passive diffusion.

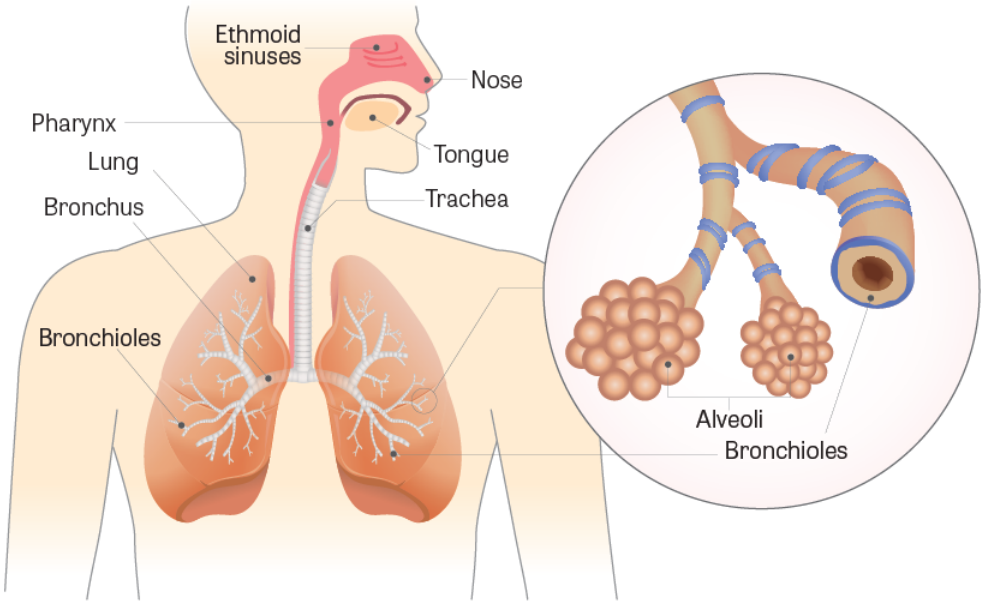


Figure 2.6.6.10.3.2. Detailed view of alveoli and bronchioles
(Image Source: adapted from iStock Photos, ©)

Impact of Physical Form on Absorption

In addition to solubility, the ability to be absorbed depends highly on the physical form of the agent (that is, whether the agent is a gas/vapor or a particle). The physical form determines the extent of its penetration into the deep lung.

Gases and Vapors

A **gas** or **vapor** can be inhaled deep into the lung and if it has high solubility in the blood, it is almost completely absorbed in one respiration (a single breath). Absorption through the alveolar membrane is by passive diffusion, following the concentration gradient. As the agent dissolves in the circulating blood, it leaves the lung and a large amount of gas or vapor can be absorbed and enter the body.

Blood-soluble gases or vapors can often be exhaled instead of absorbed right away. For blood-soluble gases, the equilibrium between the concentration of the agent in the inhaled air and that in the blood is difficult to achieve. Inhaled gases or vapors, which have poor solubility in the blood, have a limited capacity for absorption. The main reason for this is that the blood can become quickly saturated. Once saturated, blood will not be able to accept the gas and it will remain in the inhaled air and then get exhaled. One way that the amount of gas absorbed could increase is if the rate and depth of breathing were increased (this concept is known as ventilation limitation). More specifically, the amount of gas absorbed could be increased if the rate of blood supply to the lung were increased by **flow limitation**.

In contrast, *insoluble* gases or vapors can be absorbed into the body by the lungs before getting exhaled (for example, nitrogen dioxide and carbon monoxide). This is because the equilibrium between the inhaled air and the blood is reached more quickly for relatively insoluble gases than for soluble gases.

Airborne Particles

The absorption of airborne particles is usually quite different from that of gases or vapors. The absorption of solid particles, regardless of solubility, depends upon particle size:

- Large particles (>5 μM) are generally deposited in the nasopharyngeal (head airways region) region with little absorption.
- Particles 2-5 μM can penetrate into the tracheobronchial region.
- Very small particles (<1 μM) are able to penetrate deep into the alveolar sacs where they can deposit and be absorbed.

Differences in Absorption Among Regions of the Respiratory Tract

Nasopharyngeal Region

Minimal absorption takes place in the nasopharyngeal region due to the cell thickness of the mucosa and the rapid movement of gases and particles through the region.

Tracheobronchial Region

Relatively soluble gases can quickly enter the blood stream. Most deposited particles are moved back up to the mouth where they are swallowed.

Pulmonary Region

Absorption in the **alveoli** of the pulmonary region is quite efficient compared to other areas of the respiratory tract. Relatively soluble materials (gases or particles) are quickly absorbed into systemic circulation. Pulmonary macrophages exist on the surface of the alveoli. They are not fixed and not a part of the alveolar wall. They can engulf particles just as they engulf and kill microorganisms. These alveolar macrophages can scavenge and clear some insoluble particle into the lymphatic system.

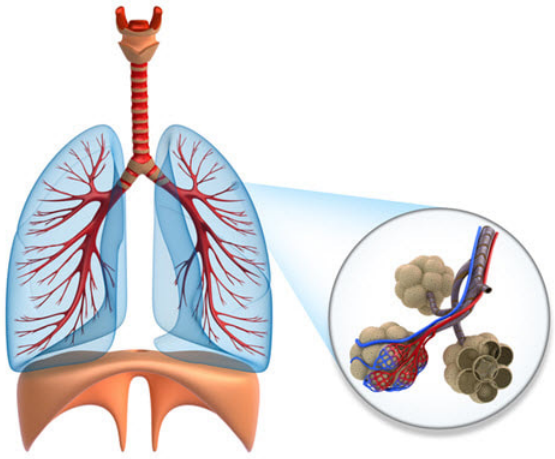


Figure 2.6.6.10.3.3. Alveoli
(Image Source: iStock Photos, ©

Some other particles may remain in the alveoli indefinitely. For example:

- Coal dust and asbestos fibers may lead to black lung or asbestosis, respectively.
- Carbon nanotubes (CNT), tiny tube-shaped structures smaller than a human hair, have been found in the lungs long after exposure. CNT are used in materials like polymers and anti-static packaging for their electric, magnetic, and mechanical properties. Studies of what happens to different forms of single-walled and multi-walled carbon nanotubes (CNT) found that pristine CNT could remain in the lung for months or even years after pulmonary deposition. However, some CNT can move to the gastrointestinal (GI) tract via the mucocilary escalator where, if swallowed, there appears to be no uptake of CNT from the GI tract (with a possible exception of the smallest functionalized single-walled CNT). In addition, under some experimental

conditions in animals, some carbon nanotubes moved from the alveolar space to the nearby pulmonary region including lymph nodes, subpleura and pleura, and smaller amounts went to distal organs including the liver, spleen, and bone marrow.

Factors Affecting the Toxicity of Inhaled Materials

The nature of toxicity of inhaled materials depends on whether the material is absorbed or remains within the alveoli and small bronchioles. If the agent is absorbed and is lipid soluble, it can rapidly distribute throughout the body, passing through the cell membranes of various organs or into fat depots. Lipid-soluble substances take a longer time to reach equilibrium than water soluble substances. Chloroform and ether are examples of lipid-soluble substances with high blood solubility.

Non-absorbed foreign material can also cause severe toxic reactions within the respiratory system. These reactions may take the form of chronic bronchitis, alveolar breakdown (emphysema), fibrotic lung disease, and even lung cancer. In some cases, the toxic particles can kill the alveolar macrophages, which results in lowering the body's respiratory defense mechanism.

Pharmaceuticals Targeted to the Respiratory Tract

Inhaled drug delivery devices can be very effective and safe for getting active agents directly to their site of action. Inhalation is used to deliver locally acting drugs to treat respiratory conditions, including asthma, chronic obstructive pulmonary disease (COPD), and airway infections. Advantages of targeted delivery to the lungs include a more rapid onset of action and an increased therapeutic effect. Depending on the drug inhaled, there can be reduced systemic side effects since a lower dose can deliver the required local concentration.

Toxicogenomics and Toxicity Testing

Toxicogenomics applies genomics concepts and technologies to study gene and protein activities within a type of tissue or type of cell in response to a chemical exposure. Toxicogenomics helps in the understanding of what genes and proteins interact with a chemical, and what diseases are associated with various genes, proteins, and chemicals. This example used toxicogenomics to evaluate the response of the respiratory tract to one type of inhaled material:
For example, titanium dioxide nanoparticles (TiO2NPs) induce lung inflammation in experimental animals and **one study** included a comprehensive toxicogenomic analysis of lung responses in mice exposed to six individual TiO2NPs exhibiting different sizes, crystalline structure, and surface modifications. The goal was to investigate whether the mechanisms leading to TiO2NP-induced lung inflammation are property specific. The results suggest that the severity of lung inflammation is property specific; however, the underlying mechanisms (genes and pathways perturbed) leading to inflammation were the same for all particle types. While the particle size clearly influenced the overall acute lung responses, a combination of small size, crystalline structure, and hydrophilic surface contributed to the long-term pathological effects observed at the highest dose.

Knowledge Check

1) An inhaled material will most likely be absorbed into the body if it has the following characteristics:

a) High lipid solubility and poorly ionized
b) Large particle size and low water solubility
c) High water solubility and small particle size

Answer

High water solubility and small particle size - **This is the correct answer.**
In contrast to absorption via the gastrointestinal tract or through the skin, gases and particles, which are water-soluble (and

thus blood soluble), will be absorbed more efficiently from the lung alveoli. Very small particles (<1 µM) are able to penetrate deep into the alveolar sacs where they can deposit and be absorbed.

2) Particles of size 2-5 µM are most likely to settle out in which location of the respiratory tract?

a) Nasopharyngeal region
b) Tracheobronchial region
c) Pulmonary region

Answer

Tracheobronchial region - **This is the correct answer.**
Particles 2-5 µM can penetrate into the tracheobronchial region. Very small particles (<1 µM) are able to penetrate deep into the alveolar sacs where they can deposit and be absorbed.

This page titled 2.6.6.10.3: Respiratory Tract is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **10.3: Respiratory Tract** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Dermal Route

In contrast to the thin membranes of the respiratory alveoli and the gastrointestinal villi, the skin is a complex, multilayer tissue. It is relatively impermeable to most ions and aqueous solutions, and serves as a barrier to most xenobiotics.

Did you know?

Dimethyl sulfoxide (DMSO) has been used in research, human and veterinary medicine, and as a solvent. After applying to the skin, some people can quickly detect a garlic taste as the DMSO is absorbed and enters the body. DMSO also increases the rate of absorption of some other compounds through the skin.

For transdermal drug delivery (TDD), the big challenge is the barrier property of skin, especially the stratum corneum (SC). Different methods have been developed to enhance the penetration of drugs through the skin, with the most popular approach being the use of penetration enhancers (PEs), including natural terpenes. Terpenes, a large and diverse class of organic compounds produced by a variety of plants, are a very safe and effective class of PEs. Limonene is one example of a terpene used as a penetration enhancer. The main mechanism for the penetration enhancing action of terpenes is the interaction with SC intercellular lipids. The key factor affecting the enhancement is the lipophilicity of the terpenes and the drug molecules.



Entry of Toxicants via Skin

Some notable toxicants can gain entry into the body following skin contamination. For example:

- Certain commonly used organophosphate pesticides have poisoned agricultural workers following dermal exposure.
- The neurological warfare agent, sarin, readily passes through the skin and can produce quick death to exposed persons.
- Several industrial solvents can cause systemic toxicity by penetrating the skin. For example:
  - Carbon tetrachloride enters the skin and causes liver injury.
  - Hexane can pass through the skin and cause nerve damage.

The skin consists of three main layers of cells as illustrated in Figure 2.6.6.10.4.1:

1. Epidermis
2. Dermis
3. Subcutaneous tissue

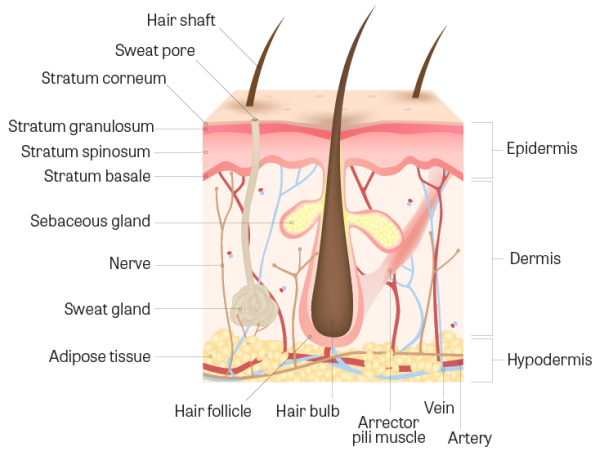


Figure 2.6.6.10.4.1. Layers of the skin
(Image Source: Adapted from iStock Photos, ©)

Epidermis and Stratum Corneum

The epidermis (and particularly the stratum corneum) is the only layer that is important in regulating the penetration of a skin contaminant. It consists of an outer layer of cells, packed with keratin, known as the **stratum corneum** layer. The stratum corneum is devoid of blood vessels. The cell walls of the keratinized cells are apparently double in thickness due to the presence of the keratin, which is chemically resistant and an impenetrable material. The blood vessels are usually about 100 μM from the skin surface. To enter a blood vessel, an agent must pass through several layers of cells that are generally resistant to penetration by chemicals.

Factors Influencing Penetration of the Stratum Corneum

Thickness

The thickness of the stratum corneum varies greatly with regions of the body. The stratum corneum of the palms and soles is very thick (400-600 μM) whereas that of the arms, back, legs, and abdomen is much thinner (8-15 μM). The stratum corneum of the axillary (underarm) and inguinal (groin) regions is the thinnest with the scrotum especially thin. As expected, the ability of toxicants to penetrate that stratum corneum inversely relates to the thickness of the epidermis.

Damage

Any process that removes or damages the stratum corneum can enhance penetration of a xenobiotic. Abrasion, scratching, or cuts to the skin will make it more penetrable. Some acids, alkalis, and corrosives can injure the stratum corneum and make it easier for agents to penetrate this layer. The most prevalent skin conditions that enhance dermal absorption are skin burns and dermatitis.

Passive Diffusion

Toxicants move across the stratum corneum by passive diffusion. There are no known active transport mechanisms functioning within the epidermis. Polar and nonpolar toxicants diffuse through the stratum corneum by different mechanisms:

- **Polar compounds,** which are water soluble, appear to diffuse through the outer surface of the hydrated keratinized layer.
- **Nonpolar compounds,** which are lipid soluble, dissolve in and diffuse through the lipid material between the keratin filaments.

Water

Water plays an important role in dermal absorption. Normally, the stratum corneum is partially hydrated (approximately 7% by weight). Penetration of polar substances is about 10 times more effective than when the skin is completely dry. Additional hydration on the skin's surface increases penetration by 3–5 times, which further increases the ability of a polar compound to penetrate the epidermis.

Species

Skin penetration can vary by species which can influence the selection of species used for safety testing. Penetration of chemicals through the skin of the monkey, pig, and guinea pig is often similar to that of humans. The skin of the rat and rabbit is generally more permeable whereas the skin of the cat is generally less permeable. For practical reasons and to assure adequate safety, the rat and rabbit have been used for dermal toxicity safety tests.

Other Sites of Dermal Absorption

In addition to the stratum corneum, small amounts of chemicals may be absorbed through the sweat glands, sebaceous glands, and hair follicles. However, since these structures represent only a very small percentage of the skin's total surface area, they are not ordinarily viewed as important contributors to dermal absorption.

Dermis and Subcutaneous Tissue

Once a substance penetrates through the stratum corneum, it enters lower layers of the epidermis, the dermis, and subcutaneous tissue. These layers are far less resistant to further diffusion. They contain a porous, nonselective aqueous diffusion medium which can be penetrated by simple diffusion. Most toxicants that have passed through the stratum corneum can now readily move through the remainder of the skin and enter the circulatory system via the large numbers of venous and lymphatic capillaries in the dermis.

Semivolatile Organic Compounds (SVOCs)

Exposure to **semivolatile organic compounds (SVOCs)** via the dermal route can occur. The amount of SVOCs absorbed via air-to-skin uptake has been estimated to be comparable to or larger than the amount taken in via inhalation for many SVOCs encountered indoors, including:

- Butylated hydroxytoluene (BHT)
- Chlordane
- Chlorpyrifos
- Diethyl phthalate
- Nicotine (in free-base form)
- Other chemicals

The influence of particles and dust on dermal exposure, the role of clothing and bedding as transport vectors, and the potential significance of hair follicles as transport shunts through the epidermis are all areas of research interest.

Human exposure to indoor SVOCs through the dermal pathway has often been underestimated and not considered in exposure assessments. However, exposure scientists, risk assessors, and public health officials are increasingly aware of and interested in the health impacts of dermal exposure. Further, experts seek to understand how health consequences can vary by the exposure pathway. For example, an SVOC that enters the blood through the skin does not encounter the same detoxification pathways that it would encounter when ingested and processed by the stomach, intestines, and liver before entering the blood; its direct entry into the blood can make it potentially more toxic.



Figure 2.6.6.10.4.2. Examples of SVOCs from consumer goods
(Image Source: Adapted from iStock Photos, ©)

Knowledge Check

1) The main barrier to dermal absorption is the:
a) Stratum corneum

b) Dermis
c) Subcutaneous tissue

Answer

Stratum corneum - **This is the correct answer.**
The epidermis (and particularly the stratum corneum) is the only layer that is important in regulating penetration of a skin contaminant.

2) The two primary factors that can increase dermal penetration are:

a) Neutralizing pH and aerosolizing
b) Increasing hydration and disruption of the stratum corneum
c) Dehydrating a substance and increasing particle size

Answer

Increasing hydration and disruption of the stratum corneum - **This is the correct answer.**
Water plays an important role in dermal absorption. Normally, the stratum corneum is partially hydrated (~7% by weight). Penetration of polar substances is about 10 times as effective as when the skin is completely dry. Additional hydration can increase penetration by 3-5 times which further increases the ability of a polar compound to penetrate the epidermis. Any process that removes or damages the stratum corneum can enhance penetration of a xenobiotic.

This page titled 2.6.6.10.4: Dermal Route is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **10.4: Dermal Route** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Other Routes of Exposure

In addition to the common routes of environmental, occupational, and medical exposure (oral, respiratory, and dermal), other routes of exposure may be used for medical purposes. Many pharmaceuticals are given by parenteral routes via injection into the body using a syringe and hollow needle.

Other Exposure Routes

Intradermal injections are made directly into the skin, just under the stratum corneum. Tissue reactions are minimal and absorption is usually slow. A **subcutaneous injection** is beneath the skin. Since the subcutaneous tissue is quite vascular (consisting of vessels especially those carrying blood), absorption into the systemic circulation is generally rapid. Tissue sensitivity is also high and thus irritating substances may induce pain and an inflammatory reaction.

The **intramuscular** route is used to inject many pharmaceuticals, especially antibiotics and vaccines, directly into muscle tissue. It is an easy procedure and the muscle tissue is less likely to become inflamed compared to subcutaneous tissue. Absorption from muscle is about the same as from subcutaneous tissue.

The **intravenous** (vein) or **intra-arterial** (artery) routes are used to inject substances directly into large blood vessels when they are irritating or when an immediate action is desired, such as anesthesia.

Parenteral injections may also be made directly into body cavities, rarely in humans but frequently in laboratory animal studies. An **intraperitoneal injection** goes directly into the abdominal cavity and an **intrapleural injection** directly into the chest cavity. Since the pleura and peritoneum have minimal blood vessels, irritation is usually minimal and absorption is relatively slow.



Figure 2.6.6.10.5.1. Injection using a syringe
(Image Source: iStock Photos, ©)

Implantation is another route of exposure of increasing concern. A large number of pharmaceuticals and medical devices are now implanted in various areas of the body. Implants may be used to allow slow, time-release of a substance such as hormones. Implanted medical devices and materials like artificial lenses, tendons, and joints, and cosmetic reconstruction do not involve absorption.

Some materials enter the body via **skin penetration** as the result of accidents or violence (weapons, etc.). The absorption in these

cases depends highly on the nature of the substance. Metallic objects (such as bullets) may be poorly absorbed whereas more soluble materials that thrust through the skin and into the body from accidents may be absorbed rapidly into the circulation.

Novel methods of introducing substances into specific areas of the body are often used in medicine. For example, **conjunctival instillations** (eye drops) treat ocular conditions where high concentrations are needed on the outer surface of the eye, not possible by other routes.

Therapies for certain conditions require that a substance is deposited in body openings where high concentrations and slow release may be needed while keeping systemic absorption to a minimum. For these substances, the pharmaceutical agent is suspended in a poorly absorbed material such as beeswax with the material known as a **suppository**. The usual locations for use of suppositories are the rectum and vagina.

Did you know?

Cinnamic aldehyde, also called **cinnamaldehyde**, gives cinnamon its flavor and odor. It occurs naturally in the bark of cinnamon trees and other species. Cinnamic aldehyde and cinnamic alcohol are well known in the scientific literature as being associated with skin allergy in humans. Skin allergy is also called skin sensitization or allergic contact dermatitis. Cinnamic aldehyde is a more potent sensitizer than cinnamic alcohol. The skin absorption and metabolism of cinnamic aldehyde and cinnamic alcohol play an important role in the development of skin sensitization following skin exposures. Cinnamic alcohol applied to human skin is converted to cinnamic aldehyde and cinnamic acid.

Cinnamic aldehyde is a good example of how an assessment of the risk of skin sensitization can be conducted prior to the introduction of new ingredients and products into the marketplace. A published quantitative risk assessment for cinnamic aldehyde used the understanding of its chemical, cellular, and molecular properties. By estimating the exposure to cinnamic aldehyde and knowing its allergenic potency, it was possible to assess the sensitization risk of cinnamic aldehyde in different types of consumer products. This publication applied exposure-based risk assessment tools to two hypothetical products containing cinnamic aldehyde. The risk assessment predicted that an eau de toilette leave-on product containing 1000 ppm or more of cinnamic aldehyde would pose an *unacceptable risk* of skin sensitization getting induced. However, a shampoo containing the same level of cinnamic aldehyde would pose an *acceptable risk* of skin sensitization getting induced, based on there being limited exposure to the ingredient from a rinse-off product application.



Figure 2.6.6.10.5.2. Cinnamon sticks and powder (Image Source: iStock Photos, ©)

Knowledge Check

1) If an immediate therapeutic effect is needed, the route of exposure that would most likely be used is the:

a) Intradermal route

b) Intramuscular injection
c) Intravenous injection

Answer

Intravenous injection - **This is the correct answer.**
Substances injected into the circulatory system go directly to the target tissue where immediate reactions can occur.

2) A pharmaceutical may be implanted in the body to:

a) Allow slow-release over a long period of time
b) Assure that the substance is distributed equally throughout the body
c) Reduce irritation from the substance

Answer

Allow slow-release over a long period of time - **This is the correct answer.**
Treatment with pharmaceuticals in time-release implants is a relatively new therapeutic technique that has gained popularity for long-term chronic chemotherapy.

This page titled 2.6.6.10.5: Other Routes of Exposure is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **10.5: Other Routes of Exposure** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.11: Distribution

Learning Objectives

After completing this lesson, you will be able to:

- Explain distribution and its role in toxicokinetics.
- Describe the impact of exposure route on distribution.
- Describe three models of disposition.
- Identify structural barriers to distribution.

In this section...

Topics include:

2.6.6.11.1: Introduction to Distribution

2.6.6.11.2: Influence of Route of Exposure

2.6.6.11.3: Disposition Models

2.6.6.11.4: Structural Barriers to Distribution

2.6.6.11.5: Storage Sites

Section 11: Key Points

What We've Covered

This section made the following main points:

- Distribution is the process in which an absorbed chemical moves away from the site of absorption to other areas of the body.
- An absorbed chemical passes through cell linings of the absorbing organ (skin, lung, or gastrointestinal tract) into the interstitial fluid of that organ.
- The toxicant can leave the interstitial fluid by entering local tissue cells, blood capillaries and the blood circulatory system, or the lymphatic system.
- If the toxicant gains entrance into the blood plasma, it:
  - Travels bound or unbound along with the blood.
  - May be excreted, stored, or biotransformed, or may interact or bind with cellular components.
- The volume of distribution (VD) is the total volume (in liters) of body fluids in which a toxicant is distributed.
- The route of exposure is an important factor affecting the concentration of the toxicant or its metabolites at any specific location within the blood or lymph.
  - Toxicants entering from the GI tract or peritoneum are immediately subject to biotransformation or excretion by the liver and elimination by the lung (this is often called the "first-pass effect").

- Toxicants absorbed through the lung or skin enter the blood and go directly to the heart and systemic circulation, thus being distributed to various organs before going to the liver (not subject to the first-pass effect).
- Toxicants that enter the lymph will not go to the liver first, but will slowly enter systemic circulation.
- The blood level of a toxicant depends on the site of absorption and the rate of biotransformation and excretion.

- Disposition is the combined processes of distirbution, biotransformation, and elimination. Disposition models can be:
  - One-Compartment Open Model — disposition of a substance introduced and distributed instantaneously and evenly in the body and eliminated proportionally to the amount left in the body ("first-order" rate).
  - Two-Compartment Open Model — the chemical enters and distributes in the first compartment (usually blood), then distributed to another compartment where it can be eliminated or may return to the first compartment.
    - The biological half-life, the most commonly used measure of the kinetic behavior of a xenobiotic, is the half-life for a chemical in a two-compartment model.
  - Multiple Compartment Model — the chemical involves several peripheral body compartments, including long-term storage, or biotransformation and elimination at varying rates as blood levels change.
- Organs or tissues differ in the amount of a chemical they may receive, depending on:
  - Volume of blood — organs that receive larger blood volumes potentially accumulate more of a given toxicant.
  - Tissue affinity — some tissues have a higher affinity for specific chemicals, accumulating a toxicant in great concentrations despite a rather low flow of blood.
- Structural barriers to distribution include the blood-brain barrier and the placental barrier.
- Toxicants can also be stored:
  - When bound to plasma proteins in the blood
  - In adipose tissues
  - In bone
  - In the liver
  - In the kidneys

This page titled 2.6.6.11: Distribution is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

Introduction to Distribution

So far, we have described the absorption of substances into the body. Now we will focus on what happens next to substances in the body after they are absorbed.

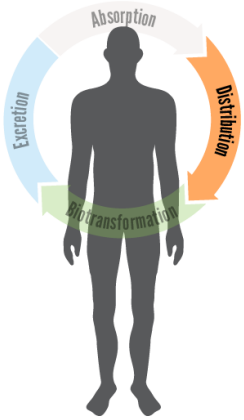


Figure 2.6.6.11.1.1. The processes of toxicokinetics
(Image Source: Adapted from iStock Photos, ©)

Distribution Defined

Distribution is the process in which an absorbed chemical moves away from the site of absorption to other areas of the body. In this section, we will answer the following questions:

- How do chemicals move through the body?
- Does distribution vary with the route of exposure?
- Is a chemical distributed evenly to all organs or tissues?
- How fast is a chemical distributed?
- Why do some chemicals stay in the body for a long time while others are eliminated quickly?

Body Fluids

When a chemical is absorbed, it passes through cell linings of the absorbing organ (skin, lung, or gastrointestinal tract) into the **interstitial fluid** (fluid surrounding cells) of that organ.
The other body fluids are **intracellular fluid** (fluid inside cells) and **blood plasma**. However, the body fluids are not isolated but represent one large pool. The interstitial and intracellular fluids, in contrast to fast moving blood, remain in place with certain components (for example, water and electrolytes) moving slowly into and out of cells. A chemical, while immersed in the interstitial fluid, is not mechanically transported the way it is in blood. Table 1 lists the approximate percentage of body weight each of these body fluids comprise.

Body fluid	Approximate % of total body weight
Interstitial fluid	15%
Intracellular fluid	40%
Blood plasma	8%

Table 1. Approximate percentages that various body fluids contribute to the total body weight

A toxicant can leave the interstitial fluid by entering:

- Local tissue cells.
- Blood capillaries and the blood circulatory system.
- The lymphatic system.

Blood Plasma

If the toxicant gains entrance into the blood plasma, it travels along with the blood, either in a bound or unbound form. Blood moves rapidly through the body via the cardiovascular circulatory system. In contrast, lymph (fluid) moves slowly through the lymphatic system. The major distribution of an absorbed chemical is by blood with only minor distribution by lymph. Since virtually all tissues have a blood supply, all organs and tissues of the body are potentially exposed to the absorbed chemical.

Distribution of a chemical to body cells and tissues requires that the toxicant penetrate a series of cell membranes. It must first penetrate the cells of the capillaries (small blood vessels) and later the cells of the target organs. The factors pertaining to passage across membranes apply to these other cell membranes as well. For example, important factors include the concentration gradient; molecular weight; lipid solubility; and polarity, with the smaller, nonpolar toxicants in high concentrations being most likely to gain entrance.

The distribution of a xenobiotic can be affected by whether it binds to plasma protein. Some toxicants may bind to these plasma proteins (especially albumin), which removes the toxicant from potential cell interaction. Within the circulating blood, the non-bound (free) portion is in equilibrium with the bound portion. However, only the free substance is available to pass through the capillary membranes. Those substances that are extensively bound are limited in terms of equilibrium and distribution throughout the body. Protein binding in the plasma greatly affects distribution, prolongs the half-life within the body, and affects the dose threshold for toxicity.

The plasma level of a xenobiotic is important since it generally reflects the concentration of the toxicant at the site of action. The passive diffusion of the toxicant into or out of these body fluids will be determined mainly by the toxicant's concentration gradient.

Volume of Distribution (written as V subscript D - see example in formula below)

The **apparent volume of distribution (VD)** is the total volume of body fluids in which a toxicant is distributed. The VD is expressed in liters.

If a toxicant is distributed only in the plasma fluid, the plasma concentration will remain high and a low VD results; however, if a toxicant is distributed in all sites (blood plasma, interstitial, and intracellular fluids) there is greater dilution in plasma concentration and a higher VD will result. Binding in effect reduces the concentration of free toxicants in the plasma or VD. Toxicants that undergo rapid storage, biotransformation, or elimination further affect the VD. Toxicologists determine the VD of a toxicant in order to know how extensively a toxicant is distributed in the body fluids. The volume of distribution can be calculated by the formula:

$$V_D = \frac{\text{dose } (mg)}{\text{plasma conc } (mg/L)}$$

The volume of distribution may provide useful estimates as to how extensive the toxicant is distributed in the body. For example, a very high apparent VD may indicate that the toxicant has distributed to a particular tissue or storage area such as adipose tissue. In addition, the body burden for a toxicant can be estimated from knowledge of the VD by using the formula:

$$\text{body burden } (mg) = \text{plasma conc } (mg/L) \times V_D\ (L)$$

Once a chemical is in the blood stream:

- It may be excreted.
- It may be stored.
- It may be biotransformed into different chemicals (metabolites).
- Its metabolites may be excreted or stored.
- The chemical or its metabolites may interact or bind with cellular components.

Most chemicals undergo some biotransformation. The degree with which various chemicals are biotransformed and the degree with which the parent chemical and its metabolites are stored or excreted vary with the nature of the exposure (dose level, frequency, and route of exposure).

Knowledge Check

1) When an ingested toxicant is absorbed, it passes through the cells lining the GI tract into the:

a) Intracellular fluid
b) Gastric fluid
c) Interstitial fluid

Answer

Interstitial fluid - **This is the correct answer.**
When a chemical is absorbed it passes through cell linings of the absorbing organ (in this case, the gastrointestinal tract) into the interstitial fluid (fluid surrounding cells) within that organ.

2) The apparent volume of distribution (VD) represents the:

a) Total volume of body fluids in which a toxicant is distributed
b) Amount of blood plasma in which a toxicant is dissolved
c) Amount of interstitial fluid that contains a toxicant

Answer

Total volume of body fluids in which a toxicant is distributed - **This is the correct answer.**
The apparent volume of distribution (VD) represents the total volume of body fluids in which a toxicant is distributed. It consists of the interstitial fluid, intracellular fluid, and the blood plasma. Soon after absorption, a toxicant may be distributed to all three types of fluids, although the concentrations may be quite different. Rarely will a toxicant be distributed to only one type of fluid.

This page titled 2.6.6.11.1: Introduction to Distribution is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **11.1: Introduction to Distribution** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.11.2: Influence of Route of Exposure

Influence of Exposure Route

The **route of exposure** is an important factor that can affect the concentration of the toxicant (or its metabolites) at any specific location within the blood or lymph. This can be important since the time and path taken by the chemical as it moves through the body influences the degree of biotransformation, storage, and elimination (and thus toxicity).

For example, if a chemical goes to the liver before going to other parts of the body, much of it may be biotransformed quickly. In this case, the blood levels of the toxicant "downstream" may be diminished or eliminated. This way of processing the chemical right away can dramatically affect its potential toxicity.

Gastrointestinal Tract and Peritoneum

When toxicants are absorbed through the **gastrointestinal (GI) tract**, a similar biotransformation process occurs. Blood carries absorbed toxicants entering the vascular system of the GI tract directly to the liver via the portal system. This is also true for those drugs administered by intraperitoneal injection. Blood from most of the peritoneum also enters the portal system and goes immediately to the liver. Blood from the liver then flows to the heart and then on to the lung, before going to other organs.

Thus, toxicants entering from the GI tract or peritoneum are immediately subject to biotransformation or excretion by the liver and elimination by the lung. This is often referred to as the "first-pass effect."

For example, first-pass biotransformation of the drug propranolol (cardiac depressant) is about 70% when given orally. This means the blood level of this medication is only about 30% of a comparable dose administered intravenously.

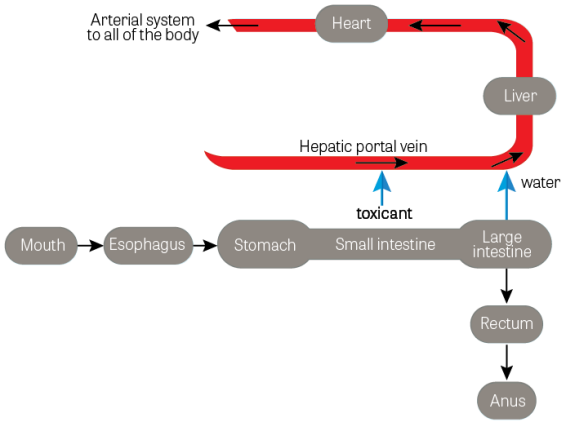


Figure 2.6.6.11.2.1. Movement of a toxicant through the portal system
(Image Source: Adapted from Kimball's Biology Pages. Original author: John W. Kimball, obtained under Creative Commons Attribution 3.0 Unported license, ©. View original image.)

Lung and Skin

Drugs and other substances that are absorbed through the **lungs** or **skin** enter the bloodstream to be carried throughout the body. Thus, they avoid the liver (hepatic) first-pass effect that would have occurred if they had been absorbed from the gastrointestinal tract. These substances can have local effects in the lungs or skin in addition to having systemic effects, and some cells in the lungs and skin may metabolize the drug or other substance. Examples of a "local first-pass effect" in the skin due to metabolism are when nitroglycerin and cortisol applied to the skin. Drugs administered intravenously or intramuscularly also enter the bloodstream to be carried throughout the body and avoid the liver (hepatic) first-pass effect.

Did you know?

Some advantages of transdermal drug delivery (skin patches):

- They are a better way to deliver substances that are broken down by the stomach acids, not well absorbed from the gut, or extensively broken down by the liver.
- They are a substitute for oral route.
- They permit constant dosing rather than the peaks and valley in medication level associated with orally administered medication.
- They can minimize undesirable side effects.
- They can be used to prescribe drugs that have short biological half-lives or a narrow therapeutic window.
- They can be removed, thereby terminating therapy easily.
- They are noninvasive, avoiding the inconvenience of IV therapy or injections.
- They can be used with patients who are nauseated or unconscious.
- They are cost-effective.



Figure 2.6.6.11.2.2. Nicotine patch
(Image Source: iStock Photos, ©)

Lymph

The delivery of drugs and bioactive compounds via the lymphatic system avoids first-pass metabolism by the liver and increases oral bioavailability. It is also a way to deliver drugs for diseases that spread through the lymphatic system such as certain types of cancer and the human immunodeficiency virus (HIV). For example, liposomes composed of phosphatidylethanol can enhance the oral bioavailability of poorly absorbed hydrophilic drugs such as cefotaxime.

Blood

The blood levels of a drug or other substance depend on the site of absorption, whether being absorbed after subcutaneous injection or more quickly from intramuscular injection. These blood levels also depend on the individual's rate of local and systemic biotransformation, and the rate of excretion. Uptake and release can occur in areas of the body away from the first site of absorption. Some anesthetics can be taken up by the lungs and later released, impacting blood levels. Lidocaine, given intravenously, is one example of this later release. Further, as noted elsewhere in ToxTutor, the metabolism of a substance can vary widely from person-to-person due to factors such as genetic differences, age, diet, and diseases that affect metabolism.

Some advantages of intramuscular injections:

- They are absorbed faster than subcutaneous injection, partly because muscle tissue has a larger blood supply than tissue just under the skin.
- They can hold a greater injected volume of drug (or vaccine) than a subcutaneous tissue injection can.
- They can be used instead of intravenous injection if a drug is irritating to veins or if a suitable vein cannot be located.
- They may be used instead of oral delivery if a drug is known to be degraded by stomach acids.

Knowledge Check

1) The main difference in distribution of a toxicant absorbed from the gastrointestinal tract from toxicants absorbed through the skin or from inhalation is:

a) The toxicant is distributed to more organs
b) A greater amount of the toxicant that is absorbed will be distributed to distant parts of the body
c) The toxicant enters the systemic circulatory system after first passing through the liver

Answer

The toxicant enters the systemic circulatory system after first passing through the liver - **This is the correct answer.**
Toxicants that enter the vascular system of the gastrointestinal tract are carried directly to the liver by the portal system. Thus, toxicants are immediately subject to biotransformation or excretion by the liver. This is often referred to as the "first pass effect."

This page titled 2.6.6.11.2: Influence of Route of Exposure is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **11.2: Influence of Route of Exposure** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.11.3: Disposition Models

Disposition Models

Disposition is the term often used to describe the combined processes of distribution, biotransformation, and elimination. The most used disposition models are compartmental models, which are categorized as one-compartment, two-compartment, and multicompartment models. Compartmental models can be used to predict the time course of drug concentrations in the body. These compartments could represent a group of similar tissues or fluids.
For example:

- Blood is a compartment.
- Fat (adipose) tissue, bone, liver, kidneys, and brain are other major compartments.

Kinetic models may be a 1) one-compartment open model; 2) two-compartment open model; or 3) multiple compartment model.

One-Compartment Model

A **one-compartment open model** may be used for drugs like aminoglycosides which rapidly distribute (equilibrate) to tissues and fluids within the body. In other words, the entire body acts like one uniform compartment. This model is also called a one-compartment open model, with "open" being the assumption that the drug can enter and leave the body via excretion. The figure below shows the disposition of a drug or other substance that distributes instantaneously and evenly in the body, and is eliminated at a rate and amount that is proportional to the amount left in the body. This is known as a "first-order" rate and represented as the logarithm of concentration in blood as a linear function of time (Figure 2.6.6.11.3.1).

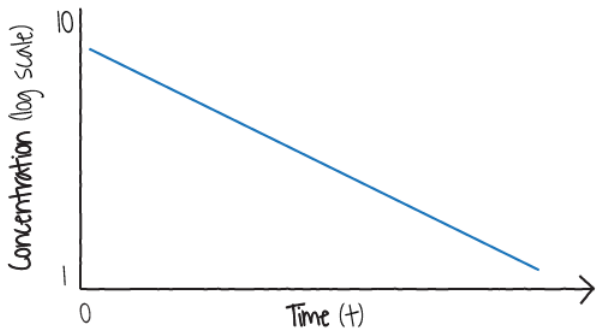


Figure 2.6.6.11.3.1. One-compartment open model
(Image Source: NLM)

The **half-life** of the chemical that follows a one-compartment model is simply the time required for half of the chemical to be lost from the plasma. Only a few chemicals actually follow the simple, first-order, one compartment model.

Two-Compartment Model

For most chemicals, it is necessary to describe the kinetics in terms of at least a two-compartment model. A two-compartment model is used for drugs which distribute slowly within the body. This model is also called a two-compartment open model, with "open" being the assumption that the drug can enter and leave the body.
For example, a one-time (bolus) intravenous administration over a short time period could lead to a drug distributing rapidly in the blood and also to highly perfused (by blood) organs like the liver and kidneys. This would be one compartment of the two-compartment model. There would be a slower distribution to other parts of the body as the second compartment.

- Two examples are vancomycin and digoxin. As shown in Figure 2.6.6.11.3.2, the drug or other substance enters and distributes in the first compartment. It is then distributed to another compartment. The concentration in the first compartment declines with

time while the concentration in the second compartment rises, peaks, and declines as the chemical is eliminated from the body.

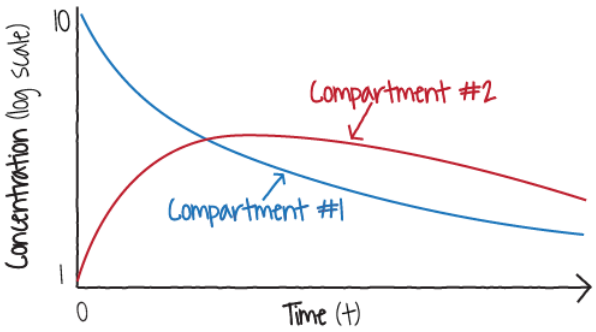


Figure 2.6.6.11.3.1. Two-compartment open model
(Image Source: NLM)

A half-life for a chemical whose kinetic behavior fits a two-compartment model is often referred to as the "biological half-life." This is the most commonly used measure of the kinetic behavior of a xenobiotic.

Multiple Compartment Model

Frequently the one- and two-compartment models cannot adequately describe the kinetics of a chemical within the body since there may be several peripheral body compartments to which the chemical may go, including long-term storage. In addition, biotransformation and elimination of a chemical may not be simple processes but subject to different rates as blood levels change.

Knowledge Check

1) Disposition models describe:

a) How a toxicant moves within the body over time
b) How the body eliminates the toxicant
c) The pathway for biotransformation of the toxicant within the liver

Answer

How a toxicant moves within the body over time - **This is the correct answer.**
Disposition models, also known as kinetic models, describe how a toxicant moves within the body compartments with time.

This page titled 2.6.6.11.3: Disposition Models is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **11.3: Disposition Models** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Structural Barriers to Distribution

Organs or tissues differ in the amount of a chemical that they receive or to which they are exposed. This is primarily due to two factors: the 1) **volume of blood** flowing through a specific tissue and the 2) **presence of special barriers** to slow down a toxicant's entrance.

Volume of Blood and Tissue Affinity

Organs that receive larger **blood volumes** can potentially accumulate more of a given toxicant. Body regions that receive a large percentage of the total cardiac output include the liver (28%), kidneys (23%), heart muscle, and brain. Bone and adipose tissues have relatively low blood flow, even though they serve as primary storage sites for many toxicants. This is especially true for toxicants that are fat-soluble and those that readily associate (or form complexes) with minerals commonly found in bone.

Tissue affinity determines the degree of concentration of a toxicant. In fact, some tissues have a higher affinity for specific chemicals and accumulate a toxicant in great concentrations despite a rather low flow of blood.

For example, adipose tissue, which has a meager blood supply, concentrates lipid-soluble toxicants. Once deposited in these storage tissues, toxicants may remain for long periods, due to their solubility in the tissue and the relatively low blood flow.

Structural Barriers

During distribution, the passage of toxicants from capillaries into various tissues or organs is not uniform. **Structural barriers** exist that restrict the entrance of toxicants into certain organs or tissues. The primary barriers are those of the brain, placenta, and testes.

Blood-Brain Barrier

The **blood-brain barrier** protects the brain from most toxicants. Specialized cells called astrocytes possess many small branches, which form a barrier between the capillary endothelium and the neurons of the brain. Lipids in the astrocyte cell walls and very tight junctions between adjacent endothelial cells limit the passage of water-soluble molecules. The blood-brain barrier is not completely impenetrable and its penetrability can vary with health status/disease state, but it does slow down the rate at which toxicants cross into brain tissue while allowing essential nutrients, including oxygen, to pass through.

Placental Barrier

The **placental barrier** protects the sensitive, developing fetus from most toxicants distributed in the maternal circulation. This barrier consists of several cell layers between the maternal and fetal circulatory vessels in the placenta. Lipids in the cell membranes limit the diffusion of water-soluble toxicants. However, nutrients, gases, and wastes of the developing fetus can pass through the placental barrier. As in the case of the blood-brain barrier, the placental barrier is not completely impenetrable but effectively slows down the diffusion of most toxicants from the mother into the fetus.

Knowledge Check

1) Organs may differ greatly in the concentration of a toxicant in them, due primarily to the:

a) Rate of elimination of the toxicant by the kidneys

b) Distance of the organ from the heart since the toxicant disintegrates quickly in the blood plasma

c) Volume of blood flow and the presence of special barriers

Answer

Volume of blood flow and the presence of special barriers - **This is the correct answer.**

Organs or tissues differ in the amount of a chemical that they receive or to which they are exposed. This is primarily due to two factors, the **volume of blood** flowing through a specific tissue and the **presence of special "barriers"** to slow down toxicant entrance. Organs that receive larger blood volumes can potentially accumulate more of a given toxicant.

2) The placental barrier protects the fetus from toxicants in the maternal blood because:

a) Substances in the maternal blood must move through several layers of cells in order to gain entrance to placental blood

b) The placenta does not contain circulating fetal blood that can absorb toxicants from the maternal blood

c) Toxicants in maternal blood are usually lipid soluble and must be water-soluble in order to penetrate through the placental cell layers

Answer

Substances in the maternal blood must move through several layers of cells in order to gain entrance to placental blood - **This is the correct answer.**

The placental barrier protects the developing and sensitive fetus from most toxicants distributed in the maternal circulation. This barrier consists of several cell layers between the maternal and fetal circulatory vessels in the placenta. Lipids in the cell membranes limit the diffusion of water-soluble toxicants.

This page titled 2.6.6.11.4: Structural Barriers to Distribution is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **11.4: Structural Barriers to Distribution** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Storage Sites

Storage of toxicants in body tissues sometimes occurs. Initially, when a toxicant enters the blood plasma, it may be bound to plasma proteins. Toxicants attached to proteins are considered a form of storage because they do not contribute to the chemical's toxic potential. Albumin is the most abundant plasma protein that binds toxicants. Normally, the toxicant is only bound to the albumin for a relatively short time.

The primary sites for toxicant storage are adipose tissue, bone, liver, and kidneys.

Adipose Tissue

Lipid-soluble toxicants are often stored in **adipose tissues**. Adipose tissue is located in several areas of the body but mainly in subcutaneous tissue. Lipid-soluble toxicants can be deposited along with triglycerides in adipose tissues. The lipids are in a continual exchange with blood and thus the toxicant may be mobilized into the blood for further distribution and elimination, or redeposited in other adipose tissue cells.

Bone

Bone is another major site for storage. Bone is composed of proteins and the mineral salt hydroxyapatite. Bone contains a sparse blood supply but is a live organ. During the normal processes that form bone, calcium and hydroxyl ions are incorporated into the hydroxyapatite-calcium matrix. Several chemicals, primarily elements, follow the same kinetics as calcium and hydroxyl ions and therefore can be substituted for them in the bone matrix.

For example, strontium (Sr) or lead (Pb) may be substituted for calcium (Ca), and fluoride (F-) may be substituted for hydroxyl (OH-) ions. Bone is continually being remodeled under normal conditions. Calcium and other minerals are continually being resorbed and replaced, on the average about every 10 years. Thus, any toxicants stored in the matrix will eventually be released to re-enter the circulatory system.

Liver and Kidneys

The **liver** is a storage site for some toxicants. It has a large blood flow and its hepatocytes (that is, liver cells) contain proteins that bind to some chemicals, including toxicants.
As with the liver, the **kidneys** have a high blood flow, which preferentially exposes these organs to toxicants in high concentrations. Storage in the kidneys is associated primarily with the cells of the nephron (the functional unit for urine formation).

Knowledge Check

1) The areas of the body which most frequently store toxicants are:

a) Adrenal gland, thyroid gland, and pancreas
b) Adipose tissue, bone, liver, and kidney
c) Skeletal muscle, tendons, and leg joints

Answer

1) Adipose tissue, bone, liver, and kidney - **This is the correct answer.**
The primary sites for toxicant storage are adipose tissue, bone, liver and kidneys. Lipid-soluble toxicants store in adipose tissues; chemicals that follow calcium or hydroxyl ion kinetics store in bone; and the liver and kidney cells are subjected to high concentrations of toxicants.

This page titled 2.6.6.11.5: Storage Sites is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **11.5: Storage Sites** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.12: Biotransformation

Learning Objectives

After completing this lesson, you will be able to:

- Explain biotransformation, including its importance to survival and the body sites it involves.
- Define enzymes and the three types of enzyme specificity.
- Explain the two phases of biotransformation.
- Identify factors that influence the effectiveness of biotransformation.

In this section...

Topics include:

2.6.6.12.1: Introduction to Biotransformation

2.6.6.12.2: Chemical Reactions

2.6.6.12.3: Biotransformation Sites

2.6.6.12.4: Modifiers of Biotransformation

Section 12: Key Points

What We've Covered

This section made the following main points:

- Biotransformation is the process by which a substance changes from one chemical to another (transformed) by a chemical reaction within the body.
- Biotransformation is vital to survival because it transforms absorbed nutrients into substances required for normal body functions.
- Potential complications of biotransformation include:
  - Detoxification — biotransformation results in metabolites of lower toxicity than the parent substance.
  - Bioactivation — biotransformation results in metabolites of greater toxicity than the parent substance.
- Chemical reactions continually occur in the body to build up new tissue, tear down old tissue, convert food to energy, dispose of waste materials, and eliminate toxic xenobiotics.
- Enzymes are catalysts for nearly all biochemical reactions in the body; essential biotransformation reactions would be slowed or prevented without these enzymes, causing major health problems.
- There are generally three types of enzyme specificity:
  1. Enzymes with absolute specificity catalyze only one reaction.
  2. Enzymes with group specificity act only on molecules that have specific functional groups.
  3. Enzymes with linkage specificity act on a particular type of chemical bond regardless of the rest of the molecular structure.
- There are two biotransformation reaction phases:
  1. Phase I reactions modify the chemical by adding a functional structure, allowing the substance to "fit" into a second (Phase II) enzyme:
     - Oxidation — the substrate loses electrons.

    - Reduction — the substrate gains electrons.
    - Hydrolysis — the addition of water splits the toxicant into two fragments or smaller molecules.
  2. Phase II reactions conjugate (join together) the modified xenobiotic with another substance. The most important Phase II reactions are:
    - Glucuronide conjugation, a high-capacity pathway — glucuronic acid is added directly to the toxicant or its Phase I metabolite, generally resulting in hydrophilic conjugates excreted by the kidney or bile.
    - Sulfate conjugation, a low-capacity pathway — decreases the toxicity of xenobiotics, resulting in highly polar sulfate conjugates readily secreted in the urine.
- Biotransformation sites are the:
  - Liver (primary site, which also makes it the most susceptible to damage by ingested toxicants).
  - Kidneys (about 10-30% of the liver's capacity).
  - Skin, intestines, testes, and placenta (low capacity).
- Biotransformation effectiveness depends on factors that can inhibit or induce enzymes and dose levels, including species, age, gender, genetic variability, nutrition, disease, exposure to other chemicals, and the dose level.

This page titled 2.6.6.12: Biotransformation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.6.12.1: Introduction to Biotransformation

Biotransformation is the process by which a substance changes from one chemical to another (transformed) by a chemical reaction within the body. **Metabolism** or **metabolic transformations** are terms frequently used for the biotransformation process. However, metabolism is sometimes not specific for the transformation process but may include other phases of toxicokinetics.

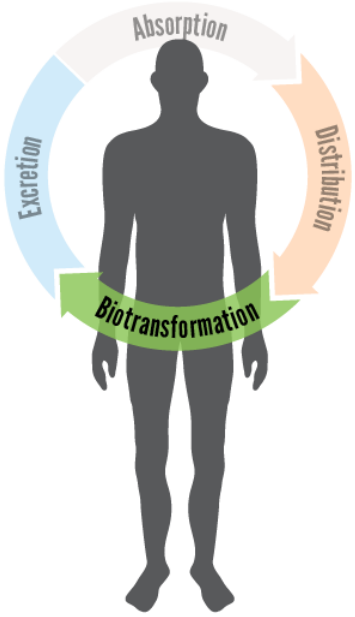


Figure 2.6.6.12.1.1. Processes of toxicokinetics (Image Source: Adapted from iStock Photos, ©)

Importance of Biotransformation

Biotransformation is vital to survival because it transforms absorbed nutrients (food, oxygen, etc.) into substances required for normal body functions. For some pharmaceuticals, it is a metabolite that is therapeutic and not the absorbed drug. For example, phenoxybenzamine, a drug given to relieve hypertension caused by pheochromocytoma, a kind of tumor, is biotransformed into a metabolite, which is the active agent. Biotransformation also serves as an important defense mechanism since toxic xenobiotics and body wastes are converted into less harmful substances and substances that can be excreted from the body.

Toxicants that are lipophilic, non-polar, and of low molecular weight are readily absorbed through the cell membranes of the skin, GI tract, and lung. These same chemical and physical properties control the distribution of a chemical throughout the body and its penetration into tissue cells. Lipophilic toxicants are hard for the body to eliminate and can accumulate to hazardous levels. However, most lipophilic toxicants can be transformed into hydrophilic metabolites that are less likely to pass through membranes of critical cells. Hydrophilic chemicals are easier for the body to eliminate than lipophilic substances. Biotransformation is thus a key body defense mechanism.

Fortunately, the human body has a well-developed capacity to biotransform most xenobiotics as well as body wastes.

Did you know?

Hemoglobin, the oxygen-carrying iron-protein complex in red blood cells, is an example of a body waste that must be eliminated. The normal destruction of aged red blood cells releases hemoglobin. Bilirubin is one of several hemoglobin metabolites. If the body cannot eliminate bilirubin via the liver because of disease, medicine, or infection, bilirubin builds up in the body and the whites of the eyes and the skin may look yellow. Bilirubin is toxic to the brain of newborns and, if present in high concentrations, may cause irreversible brain injury. Biotransformation of the lipophilic bilirubin molecule in the liver results in the production of water-soluble (hydrophilic) metabolites excreted into bile and eliminated via the feces.

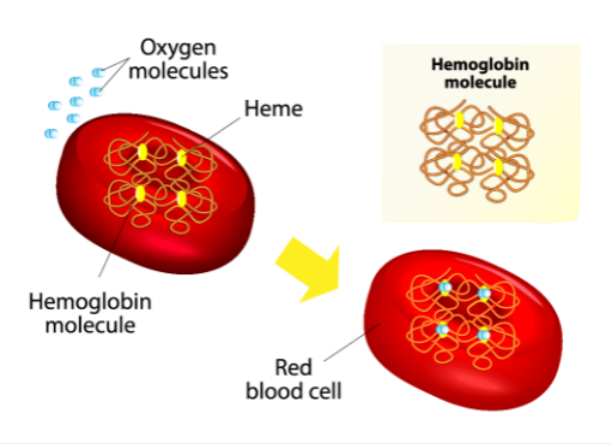


Figure 2.6.6.12.1.2. Human hemoglobin (Image Source: Adapted from iStock Photos, ©)

Potential Complications

The biotransformation process is not perfect. **Detoxification** occurs when biotransformation results in metabolites of lower toxicity. In many cases, however, the metabolites are more toxic than the parent substance, a process called **bioactivation**. Occasionally, biotransformation can produce an unusually reactive metabolite that may interact with cellular macromolecules like DNA. This can lead to very serious health effects such as cancer or birth defects.

An example is the biotransformation of vinyl chloride into vinyl chloride epoxide, which covalently binds to DNA and RNA, a step leading to cancer of the liver.

Knowledge Check

The term "biotransformation" refers to:

a. An increase in electrical charge in tissues produced by a biological transformer
b. Chemical reactions in the body that create a new chemical from another chemical
c. The transformation of one type of cell in a tissue to another type of cell

Answer

Chemical reactions in the body that create a new chemical from another chemical - **This is the correct answer.**
Biotransformation is the process whereby a substance is changed from one chemical to another (transformed) by a chemical reaction within the body.

Detoxification is a biotransformation process in which:

a. Metabolites of lower toxicity are produced
b. Metabolites of higher toxicity are produced

Answer

Metabolites of lower toxicity are produced - **This is the correct answer.**
When biotransformation results in metabolites of lower toxicity, the process is known as detoxification.

Bioactivation is a biotransformation process in which:

a. Metabolites of lower toxicity are produced
b. Metabolites of higher toxicity are produced

Answer

Metabolites of higher toxicity are produced - **This is the correct answer.**
When biotransformation results in metabolites of higher toxicity, this is known as bioactivation.

This page titled 2.6.6.12.1: Introduction to Biotransformation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **12.1: Introduction to Biotransformation** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.12.2: Chemical Reactions

Chemical Reactions

Chemical reactions are continually taking place in the body. They are a normal aspect of life, participating in the:

- Building up of new tissue.
- Tearing down of old tissue.
- Conversion of food to energy.
- Disposal of waste materials.
- Elimination of toxic xenobiotics.

Within the body is a magnificent assembly of chemical reactions, which is well orchestrated and called upon as needed. Most of these chemical reactions occur at significant rates only because specific proteins, known as enzymes, are present to catalyze them, that is, accelerate the reaction. A catalyst is a substance that can accelerate a chemical reaction of another substance without itself undergoing a permanent chemical change.

Enzymes

Enzymes are the catalysts for nearly all biochemical reactions in the body. Without these enzymes, essential biotransformation reactions would take place slowly or not at all, causing major health problems.

Did you know?

Phenylketonuria (PKU) is the genetic condition in which the enzyme that biotransforms phenylalanine to tyrosine (another amino acid) is defective. As the result, phenylalanine can build up in the body and cause severe mental retardation. Babies are routinely checked at birth for PKU. If they have PKU, they need to follow a special diet to restrict the intake of phenylalanine in infancy and childhood.



Figure 2.6.6.12.2.1. Phenylketonuia (PKU) testing in an infant
(Image Source: Wikimedia Commons, obtained under Public Domain license. Author: U.S. Air Force Photographic Archives)

These enzymatic reactions are not always simple biochemical reactions. Some enzymes require the presence of cofactors or coenzymes in addition to the **substrate** (the substance to be catalyzed) before their catalytic activity can be exerted. These co-factors exist as a normal component in most cells and are frequently involved in common reactions to convert nutrients into energy (vitamins are an example of co-factors). It is the drug or chemical transforming enzymes that hold the key to xenobiotic transformation. The relationship of substrate, enzyme, coenzyme, and transformed product can be shown as:

$$substrate * \frac{enzyme}{co-enzyme} = transformedproduct$$

Most biotransforming enzymes are high molecular weight proteins, composed of chains of amino acids linked together by peptide bonds. A wide variety of biotransforming enzymes exist. Most enzymes will catalyze the reaction of only a few substrates, meaning that they have high specificity. Specificity is a function of the enzyme's structure and its catalytic sites. While an enzyme may encounter many different chemicals, only those chemicals (substrates) that fit within the enzyme's convoluted structure and spatial arrangement will be locked on and affected. This is sometimes referred to as the "lock and key" relationship.
As shown in Figure 2.6.6.12.2.2, when a substrate fits into the enzyme's structure, an enzyme-substrate complex can be formed. This allows the enzyme to react with the substrate with the result that two different products are formed. If the substrate does not fit into the enzyme ("incompatible"), no complex will be formed and thus no reaction can occur.

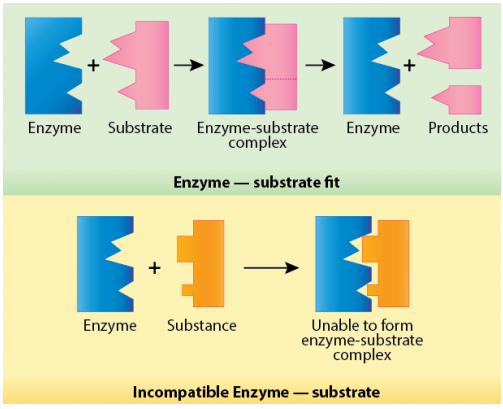


Figure 2.6.6.12.2.2. If the substrate does not fit into the enzyme, no complex will be formed and no reaction will occur.
(Image Source: NLM)

Enzyme Specificity

Enzymes range from having absolute specificity to broad and overlapping specificity. In general, there are three main types of specificity:

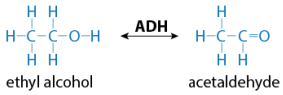
1. **Absolute** — the enzyme will catalyze only one reaction. Examples:
   - Formaldehyde dehydrogenase catalyzes only the reaction for formaldehyde.
   - Acetylcholinesterase biotransforms the neurotransmitting chemical, acetylcholine.
2. **Group** — the enzyme will act only on molecules that have specific functional groups, such as amino, phosphate, or methyl groups.
   - For example, alcohol dehydrogenase can biotransform several different alcohols, including methanol and ethanol.
3. **Linkage** — the enzyme will act on a particular type of chemical bond regardless of the rest of the molecular structure.
   - For example, N-oxidation can catalyze a reaction of a nitrogen bond, replacing the nitrogen with oxygen.

Enzyme Naming Convention

The names assigned to enzymes may seem confusing at first. However, except for some of the originally studied enzymes (such as pepsin and trypsin), a convention has been adopted to name enzymes. Enzyme names end in "ase" and usually combine the substrate acted on and the type of reaction catalyzed.

For example, alcohol dehydrogenase is an enzyme that biotransforms alcohols by the removal of a hydrogen. The result is a completely different chemical, an aldehyde or ketone.

The biotransformation of ethyl alcohol to acetaldehyde is depicted in Figure 2.6.6.12.2.3.



ADH = alcohol dehydrogenase, a specific catalyzing enzyme

Figure 2.6.6.12.2.3. Biotransformation of ethyl alcohol
(Image Source: NLM)

Beneficial or Harmful?

At this point in ToxTutor you likely see that the transformation of a specific xenobiotic can be either beneficial or harmful, and perhaps both depending on the dose and circumstances.
A good example is the biotransformation of acetaminophen (Tylenol®). When the prescribed doses are taken, the desired therapeutic response is observed with little or no toxicity. However, when excessive doses of acetaminophen are taken, hepatotoxicity can occur. This is because acetaminophen normally undergoes rapid biotransformation with the metabolites quickly eliminated in the urine and feces.
At high doses, the normal level of enzymes may be depleted and the acetaminophen is available to undergo the reaction by an additional biosynthetic pathway, which produces a reactive metabolite that is toxic to the liver. For this reason, a user of Tylenol® is warned not to take the prescribed dose more frequently than every 4–6 hours and not to consume more than four doses within a 24-hour period.
Biotransforming enzymes, like most other biochemicals, are available in a normal amount and in some situations can be "used up" at a rate that exceeds the body's ability to replenish them. This illustrates the frequently used phrase, the "dose makes the poison."



Figure 2.6.6.12.2.4. Generic acetaminophen tablets
(Image Source: iStock Photos, ©)

Biotransformation Reaction Phases

Biotransformation reactions are categorized not only by the nature of their reactions, for example, oxidation, but also by the normal sequence with which they tend to react with a xenobiotic. They are usually classified as Phase I and Phase II reactions.

Phase I reactions are generally reactions which modify the chemical by adding a functional structure. This allows the substance to "fit" into a second, or Phase II enzyme, so that it can become conjugated (joined together) with another substance.

Phase II reactions consist of those enzymatic reactions that conjugate the modified xenobiotic with another substance. The conjugated products are larger molecules than the substrate and generally polar in nature (water soluble). Thus, they can be readily excreted from the body. Conjugated compounds also have poor ability to cross cell membranes.

In some cases, the xenobiotic already has a functional group that can be conjugated and the xenobiotic can be biotransformed by a Phase II reaction without going through a Phase I reaction.

For example, phenol can be directly conjugated into a metabolite that can then be excreted. The biotransformation of benzene requires both Phase I and Phase II reactions. As illustrated in Figure 2.6.6.12.2.5, benzene is biotransformed initially to phenol by a Phase I reaction (oxidation). Phenol has a structure including a functional hydroxyl group that is then conjugated by a Phase II reaction (sulfation) to phenyl sulfate.

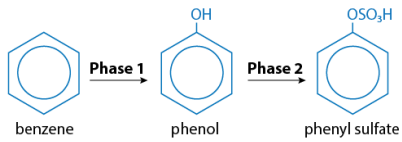


Figure 2.6.6.12.2.5. Biotransformation of benzene into phenol in Phase 1 (oxidation), which is then conjugated by a Phase 2 reaction (sulfation) to phenyl sulfate
(Image Source: NLM)

Table 1 lists the major transformation reactions for xenobiotics broken into Phase I and Phase II reactions. These reactions are discussed in more detailed below.

Phase I	Phase II
Oxidation	Sulfate conjugation
Reduction	Glucuronide conjugation
Hydrolysis	Glutathione conjugation
Acetylation	Amino acid conjugation

Table 1. Phase I and II biotransformation reactions for xenobiotics

Phase I Reactions

Phase I biotransformation reactions are simple reactions compared to Phase II reactions. In Phase I reactions, a small polar group (containing both positive and negative charges) is either exposed on the toxicant or added to the toxicant. The three main Phase I reactions are 1) oxidation; 2) reduction; and 3) hydrolysis.

Oxidation

Oxidation is a chemical reaction in which a *substrate loses electrons*. There are a number of reactions that can achieve the removal of electrons from the substrate.

- The **addition of oxygen**, or **oxygenation**, was the first of these reactions discovered and thus the reaction was named oxidation. However, many of the oxidizing reactions do not involve oxygen.
- The simplest type of oxidation reaction is **dehydrogenation**, which is the removal of hydrogen from the molecule.
- Another example of oxidation is **electron transfer** that consists simply of the transfer of an electron from the substrate.

Figure 2.6.6.12.2.6 shows these types of oxidizing reactions.

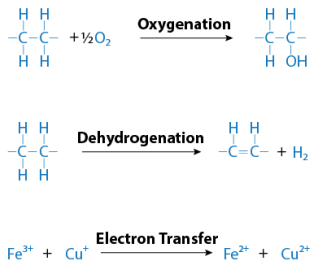


Figure 2.6.6.12.2.6. Three types of oxidation reactions
(Image Source: NLM)

The specific oxidizing reactions and oxidizing enzymes are numerous and several textbooks are devoted to this subject. Most of the reactions are described by the name of the reaction or enzyme involved. Some of these oxidizing reactions include:

- Alcohol dehydrogenation
- Aldehyde dehydrogenation
- Alkyl/acyclic hydroxylation
- Aromatic hydroxylation
- Deamination
- Desulfuration
- N-dealkylation
- N-hydroxylation
- N-oxidation
- O-dealkylation
- Sulphoxidation

Reduction

Reduction is a chemical reaction in which the *substrate gains electrons*. Reductions are most likely to occur with xenobiotics in which oxygen content is low. Reductions can occur across nitrogen-nitrogen double bonds (azo reduction) or on nitro groups (NO2). Frequently, the resulting amino compounds are oxidized which forms toxic metabolites. Some chemicals such as carbon tetrachloride can be reduced to free radicals, which are quite reactive with biological tissues. Thus, reduction reactions frequently result in activation of a xenobiotic rather than detoxification. An example of a reduction reaction in which the nitro group is reduced is illustrated in Figure 2.6.6.12.2.7.

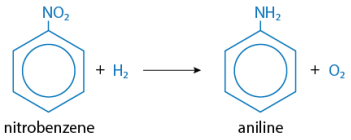


Figure above

There are fewer specific reduction reactions than oxidizing reactions. The nature of these reactions is also described by their name. Some reducing reactions include:

- Azo reduction
- Dehalogenation

- Disulfide reduction
- Nitro reduction
- N-oxide reduction
- Sulfoxide reduction

Hydrolysis

Hydrolysis is a chemical reaction in which the *addition of water* splits the toxicant into two fragments or smaller molecules. The hydroxyl group (OH-) is incorporated into one fragment and the hydrogen atom is incorporated into the other. Larger chemicals such as esters, amines, hydrazines, and carbamates are generally biotransformed by hydrolysis.
An example of hydrolysis is illustrated in the biotransformation of procaine (local anesthetic) which is hydrolyzed to two smaller chemicals (Figure 2.6.6.12.2.8).

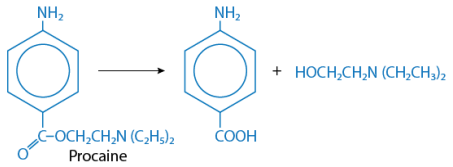


Figure 2.6.6.12.2.8. Hydrolysis of procaine
(Image Source: Adapted from Humboldt State University, Department of Chemistry. Author: Richard A. Paselk, Professor Emeritus. View original image.)

Toxicants that have undergone Phase I biotransformation are converted to metabolites that are sufficiently ionized, or hydrophilic, to be either eliminated from the body without further biotransformation or converted to an intermediate metabolite that is ready for Phase II biotransformation. The intermediates from Phase I transformations may be pharmacologically more effective and in many cases more toxic than the parent xenobiotic.

Phase II Reactions

A xenobiotic that has undergone a Phase I reaction is now a new intermediate metabolite that contains a reactive chemical group such as hydroxyl (-OH), amino (-NH2), and carboxyl (-COOH). Many of these intermediate metabolites do not possess sufficient hydrophilicity to permit elimination from the body. These metabolites must undergo additional biotransformation as a Phase II reaction.
Phase II reactions are conjugation reactions where a molecule normally present in the body is added to the reactive site of the Phase I metabolite. The result is a conjugated metabolite that is more water soluble than the original xenobiotic or Phase I metabolite. Usually, the Phase II metabolite is quite hydrophilic and can be readily eliminated from the body. The primary Phase II reactions are:

- Glucuronide conjugation – most important reaction (detailed below)
- Sulfate conjugation – important reaction (detailed below)
- Acetylation
- Amino acid conjugation
- Glutathione conjugation
- Methylation

Glucuronide Conjugation

Glucuronide conjugation is one of the most important and common Phase II reactions. The glucuronic acid molecule is used in this reaction. It is derived from glucose, a common carbohydrate (sugar) that is the primary source of energy for cells. In this reaction, glucuronic acid is added directly to the toxicant or its phase I metabolite. The sites of glucuronidation reactions are substrates having an oxygen, nitrogen, or sulfur bond, which apply to a wide array of xenobiotics as well as endogenous substances, such as bilirubin, steroid hormones, and thyroid hormones.

Glucuronidation is a pathway that conjugates xenobiotics at a high capacity ("high-capacity pathway"). Glucuronide conjugation usually decreases toxicity although there are some notable exceptions, for example, where it can result in producing carcinogenic substances. The glucuronide conjugates are generally quite hydrophilic and are excreted by the kidney or bile, depending on the size of the conjugate. The glucuronide conjugation of aniline is illustrated in Figure 2.6.6.12.2.9.

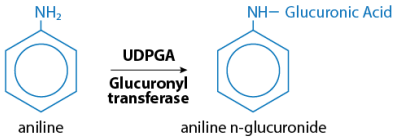


Figure 2.6.6.12.2.9. Glucuronide conjugation of aniline (which is used to make polyurethane, pharmaceuticals, and industrial chemicals)
(Image Source: NLM)

Sulfate Conjugation

Sulfate conjugation is another important Phase II reaction that occurs with many xenobiotics. In general, sulfation decreases the toxicity of xenobiotics. Unlike glucuronic acid conjugates that are often eliminated in the bile, the highly polar sulfate conjugates are readily secreted in the urine. In general, sulfation is a low-capacity pathway for xenobiotic conjugation. Often glucuronidation or sulfation can conjugate the same xenobiotics.

Knowledge Check

1) The substances in the body that accelerate chemical reactions are known as:
a) Amino acids
b) Enzymes
c) Substrates

Answer

Enzymes - **This is the correct answer.**
Enzymes are proteins that catalyze nearly all biochemical reactions in the body.

2) The convention used to name specific enzymes consists of combining:

a) The substrate name with the type of chemical reaction
b) The target organ and the type of chemical reaction
c) The substrate name with the form of toxicity

Answer

The substrate name with the type of chemical reaction - **This is the correct answer.**
Enzyme names end in "ase" and usually combine the substrate acted on and the type of reaction catalyzed.

3) Biotransformation reactions are classified as Phase I and Phase II. The basic difference is:

a) Phase I reactions conjugate a substrate whereas Phase II reactions oxidize the substance
b) Phase I reactions generally add a functional structure whereas Phase II reactions conjugate the substance
c) A Phase I reaction generally makes a substance more hydrophilic than a Phase II reaction

Answer

Phase I reactions generally add a functional structure whereas Phase II reactions conjugate the substance - **This is the correct answer.**
Phase I reactions are generally reactions which modify the chemical by adding a functional structure. This allows the substance to "fit" into the Phase II enzyme so that it can become conjugated (joined together) with another substance. Phase II reactions consist of those enzymatic reactions that conjugate the modified xenobiotic with another substance.

4) The difference between oxidation and reduction reactions is:

a) A substrate gains electrons from an oxidation reaction whereas it loses electrons by a reduction reaction
b) Oxygen is removed from a substrate in oxidation and added in the reduction reaction
c) A substrate losses electrons from an oxidation reaction whereas it gains electrons by a reduction reaction

Answer

A substrate losses electrons from an oxidation reaction whereas it gains electrons by a reduction reaction - **This is the correct answer.**
Oxidation is a chemical reaction in which a substrate loses electrons. Reduction is a chemical reaction in which the substrate gains electrons.

5) Which conjugation reaction is the most common in the biotransformation of xenobiotics?

a) Amino acid conjugation
b) Glucuronide conjugation
c) Methylation

Answer

Glucuronide conjugation - **This is the correct answer.**
Glucuronide conjugation is one of the most important and common Phase II reactions. Glucuronidation is a high-capacity pathway for xenobiotic conjugation.

This page titled 2.6.6.12.2: Chemical Reactions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **12.2: Chemical Reactions** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.12.3: Biotransformation Sites

Biotransformation Sites

Biotransforming enzymes are widely distributed throughout the body.

- The liver is the primary biotransforming organ due to its large size and high concentration of biotransforming enzymes.
- The kidneys and lungs are next with 10-30% of the liver's capacity.
- A low capacity exists in the skin, intestines, testes, and placenta.

Primary Biotransformation Site: The Liver

Since the liver is the primary site for biotransformation, it is also potentially vulnerable to the toxic action of a xenobiotic that is activated to a more toxic compound.
Within the liver cell, the primary subcellular components containing the transforming enzymes are the **microsomes (small vesicles) of the endoplasmic reticulum** and the soluble fraction of the **cytoplasm (cytosol)**. The mitochondria, nuclei, and lysosomes contain a small level of transforming activity.

- Microsomal enzymes are associated with most Phase I reactions. Glucuronidation enzymes are also contained in microsomes.
- Cytosolic enzymes are non-membrane-bound and occur free within the cytoplasm. They are generally associated with Phase II reactions, although some oxidation and reduction enzymes are contained in the cytosol.
- The most important enzyme system involved in Phase I reactions are the **cytochromes P450**, also called the cytochrome P-450 system or the mixed function oxidase (MFO) system, but now mostly called CYP450 or CYPs by scientists and in research publications. It is found in microsomes and is responsible for oxidation reactions of a wide array of chemicals.

Susceptibility of the Liver

The liver is particularly susceptible to damage by ingested toxicants because it biotransforms most xenobiotics and receives blood directly from the gastrointestinal tract. Blood leaving the gastrointestinal tract does not flow directly into the general circulatory system. Instead, it flows into the liver first via the portal vein. This process is known as the "first pass." Blood leaving the liver is eventually distributed to all other areas of the body; however, much of the absorbed xenobiotic has undergone detoxification or bioactivation. The liver may have removed most of the potentially toxic chemical. On the other hand, some toxic metabolites are highly concentrated in the liver.

Knowledge Check

1) The organ that has the greatest ability to biotransform xenobiotics is the:

a) Liver
b) Pancreas
c) Skin

Answer

Liver - **This is the correct answer.**
Biotransforming enzymes are widely distributed throughout the body. However, the liver has the largest concentration of all organs and thus has a very high capacity for biotransformation.

2) The "first pass" phenomenon pertains to:

a) The situation where xenobiotics that are absorbed from the GI tract first enter the circulating blood before going to the liver
b) A condition where the liver first biotransforms a xenobiotic by Phase II reaction before it is biotransformed by a Phase I reaction

c) An anatomical arrangement in which xenobiotics absorbed from the intestine go to the liver first rather than into the systemic circulation

Answer

An anatomical arrangement in which xenobiotics absorbed from the intestine go to the liver first rather than into the systemic circulation - **This is the correct answer.**
Blood leaving the gastrointestinal tract does not directly flow into the general circulatory system. Instead, it flows into the liver first via the portal vein. This is known as the "first pass" phenomena.

This page titled 2.6.6.12.3: Biotransformation Sites is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **12.3: Biotransformation Sites** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.12.4: Modifiers of Biotransformation

Modifiers of Biotransformation

The relative effectiveness of biotransformation depends on several factors that can inhibit or induce enzymes and dose levels. Factors include:

- Species
- Age
- Gender
- Genetic variability
- Nutrition
- Disease
- Exposure to other chemicals

Species

It is well known that the capability to biotransform specific chemicals varies by **species**. These differences are termed selective toxicity, which refers to differences in toxicity between species similarly exposed. Research uses what is known about selective toxicity to develop chemicals that are effective but relatively safe in humans.

- For example, the pesticide malathion in mammals is biotransformed by hydrolysis to relatively safe metabolites, but in insects, it is oxidized to malaoxon, which is lethal to insects.

Age and Gender

Age may affect the efficiency of biotransformation. In general, human fetuses and newborns have limited abilities to carry out xenobiotic biotransformations. This limitation is due to inherent deficiencies in many of the enzymes responsible for catalyzing Phase I and Phase II biotransformations. While the capacity for biotransformation fluctuates with age in adolescents, by early adulthood the enzyme activities have essentially stabilized. The aged also have decreased biotransformation capability.
Gender may influence the efficiency of biotransformation for specific xenobiotics. This is usually limited to hormone-related differences in the oxidizing cytochrome P-450 enzymes.

Genetic Variability

Genetic variability in biotransforming capability accounts for most of the large variation among humans. In particular, human genetic differences influence the Phase II acetylation reaction. Some persons have rapid acetylation ("rapid acetylator") while others have a slow ability to carry out this reaction ("slow acetylator"). The most serious drug-related toxicity occurs in those who have slow acetylators, often referred to as "slow metabolizers." With slow acetylators, acetylation is so slow that blood or tissue levels of certain drugs (or Phase I metabolites) exceed their toxic threshold.

Table 2.6.6.12.4.1 includes examples of drugs that build up to toxic levels in slow metabolizers who have specific genetic-related defects in biotransforming enzymes.

Drug	Drug Category	Metabolic Defect	Toxic Effect
Isoniazid	antituberculosis drug	slow acetylation	nerve and liver damage
Hydralazine	antihypertensive drug	defect in mono-oxygenase enzyme	excessive fall in blood pressure
Dapsone	antibacterial agent	slow acetylation	systemic lupus erythematosus
Primaquine	antimalarial agent	defective G6PD	acute hemolytic anemia

Table 1. Examples of drugs that build to toxic levels in slow metabolizers with specific genetic-related defacts in biotransforming enzymes

Nutrition

Poor **nutrition** can have a detrimental effect on biotransforming ability. Poor nutrition relates to inadequate levels of protein, vitamins, and essential minerals. These deficiencies can decrease a person's ability to synthesize biotransforming enzymes. Many diseases can impair an individual's capacity to biotransform xenobiotics.

For example, hepatitis (a liver disease) is well known to reduce hepatic biotransformation to less than half of its normal capacity.

Prior or Simultaneous ExposurePrior or simultaneous exposure to xenobiotics can cause enzyme inhibition and enzyme induction. In some situations, exposure to a substance will inhibit the biotransformation capacity for another chemical due to **inhibition of specific enzymes**. A major mechanism for the inhibition is competition between the two substances for the available oxidizing or conjugating enzymes. The presence of one substance uses up the enzyme needed to metabolize the second substance.

Exposure to Other Environmental Chemicals and DrugsEnzyme induction is a situation where prior exposure to certain environmental chemicals and drugs results in an enhanced capability for biotransforming a xenobiotic. The prior exposures stimulate the body to increase the production of some enzymes. This increased level of enzyme activity results in increased biotransformation of a chemical subsequently absorbed.

Examples of enzyme inducers include:

- Alcohol
- Isoniazid
- Polycyclic halogenated aromatic hydrocarbons (for example, dioxin)
- Phenobarbital
- Cigarette smoke

The most commonly induced enzyme reactions involve the cytochrome P450 enzymes.

Dose level can affect the nature of the biotransformation. In certain situations, the biotransformation may be quite different at high doses compared to low dose levels. This difference in biotransformation contributes to a dose threshold for toxicity. The existence of different biotransformation pathways can usually explain what causes this dose-related difference in biotransformation. At low doses, a xenobiotic may follow a biotransformation pathway that detoxifies the substance. However, if the amount of xenobiotic exceeds the specific enzyme capacity, the biotransformation pathway is saturated. In that case, it is possible that the level of parent toxin builds up. In other cases, the xenobiotic may enter a different biotransformation pathway that may end up producing a toxic metabolite.

An example of a dose-related difference in biotransformation occurs with acetaminophen (Tylenol®):

- At normal doses:
  - About 96% of acetaminophen is biotransformed to non-toxic metabolites by sulfate and glucuronide conjugation.
  - About 4% of the acetaminophen oxidizes to a toxic metabolite.
  - That toxic metabolite is conjugated with glutathione and excreted.
- At 7-10 times the recommended therapeutic level:
  - The sulfate and glucuronide conjugation pathways become saturated and more of the toxic metabolite is formed.

- The glutathione in the liver may also be depleted so that the toxic metabolite is not detoxified and eliminated.
- It can react with liver proteins and cause fatal liver damage.

Knowledge Check

1) Selective toxicity refers to a difference in the toxicity of a xenobiotic to different species. This selective toxicity can usually be attributed to differences in:

a) The ability to absorb the xenobiotic
b) Organ systems between species
c) Capability to biotransform the xenobiotic

Answer

Capability to biotransform the xenobiotic - **This is the correct answer.**
A difference between species in their capability to biotransform a specific chemical is normally the basis for a chemical's selective toxicity.

This page titled 2.6.6.12.4: Modifiers of Biotransformation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **12.4: Modifiers of Biotransformation** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.13: Excretion

Learning Objectives

After completing this lesson, you will be able to:

- Define excretion.
- Identify the primary organ systems involved in excretion.
- Describe the basic mechanisms of excretion within each primary organ system involved.

In this section...

Topics include:

2.6.6.13.1: Introduction to Secretion

2.6.6.13.2: Urinary Excretion

2.6.6.13.3: Fecal Excretion

2.6.6.13.4: Exhaled Air

2.6.6.13.5: Other Routes

Section 13: Key Points

What We've Covered

This section made the following main points:

- Excretion, as used in ToxTutor, pertains to the elimination of a xenobiotic and its metabolites by specific excretory organs.
- The primary organ systems involved in excretion are the:
  - Urinary system, which involves:
    1. Filtration in the glomerulus.
    2. Secretion in the proximal tubule section of the nephron to transport certain molecules out of the blood and into the urine.
    3. Reabsorption in the proximal convoluted tubule of the nephron to reenter nearly all of the water, glucose, potassium, and amino acids lost during filtration back into the blood.
  - Gastrointestinal system, which occurs from two processes:
    1. Biliary excretion — generally active secretion by the liver into the bile and then into the intestinal tract, where it can be eliminated in the feces or reabsorbed.
    2. Intestinal excretion — an important elimination route only for xenobiotics that have slow biotransformation or slow urinary or biliary excretion.
  - Respiratory system, which is important for xenobiotics and metabolites that exist in a gaseous phase in the blood:
    - Excreted by passive diffusion from the blood into the alveolus.

- Minor routes of excretion occur including breast milk, sweat, saliva, tears, and semen.

This page titled 2.6.6.13: Excretion is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

Introduction to Excretion

Elimination from the body is very important in determining the potential toxicity of a xenobiotic. When the body rapidly eliminates a toxic xenobiotic (or its metabolites), it is less likely that they will be able to concentrate in and damage critical cells. The terms excretion and elimination are frequently used to describe the same process in which a substance leaves the body. **Elimination** is sometimes used in a broader sense and includes the removal of the absorbed xenobiotic through metabolic pathways as well as through excretion. **Excretion**, as used here, pertains to the elimination of the xenobiotic and its metabolites by specific excretory organs.

Except for the lung, polar (hydrophilic) substances are more likely than lipid-soluble toxicants to be eliminated from the body. Chemicals must again pass through membranes in order to leave the body, and the same chemical and physical properties that governed passage across other membranes apply to excretory organs as well.

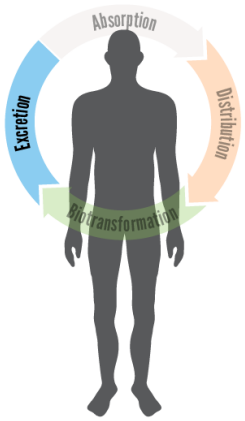


Figure 2.6.6.13.1.1. Processes of toxicokinetics
(Image Source: Adapted from iStock Photos, ©)

Primary Routes of Excretion

The body uses several routes to eliminate toxicants or their metabolites. The main routes of excretion are via urine, feces, and exhaled air. Thus, the primary organ systems involved in excretion are the:

- Urinary system
- Gastrointestinal system
- Respiratory system

A few other avenues for elimination exist but they are relatively unimportant, except in exceptional circumstances.

Knowledge Check

1) The three major routes of excretion are:

a) Gastrointestinal tract, sweat, and saliva
b) Mother's milk, tears, and semen
c) Urinary excretion, fecal excretion, and exhaled air

Answer

Urinary excretion, fecal excretion, and exhaled air - **This is the correct answer.**
The main routes of excretion are via urine, feces, and exhaled air.

This page titled 2.6.6.13.1: Introduction to Secretion is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **13.1: Introduction to Secretion** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Urinary Excretion

The primary route in which the body eliminates substances is through the kidneys. The main function of the kidney is the excretion of body wastes and harmful chemicals into the urine. The functional unit of the kidney responsible for excretion is the nephron. Each kidney contains about one million nephrons. The nephron has three primary regions that function in the renal excretion process: the glomerulus, proximal tubule, and the distal tubule (Figure 2.6.6.13.2.2).

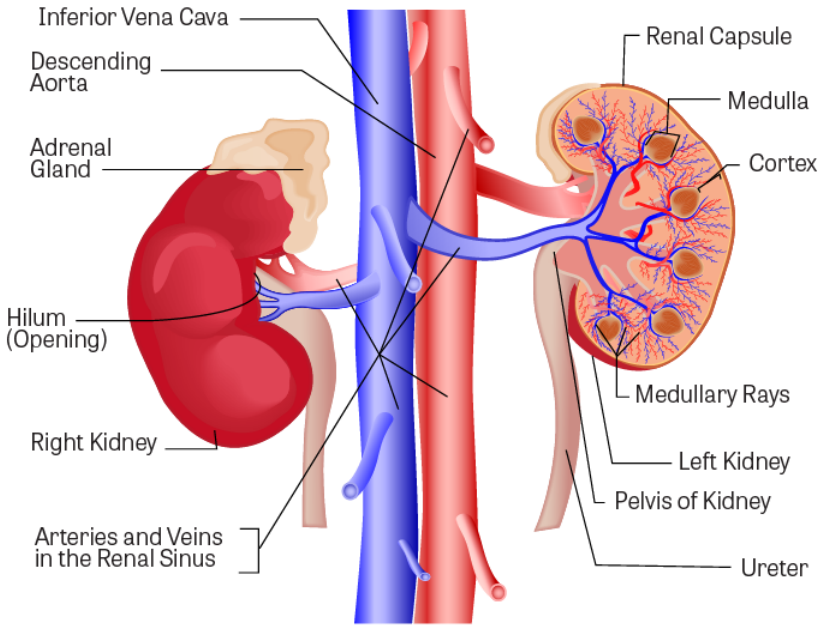
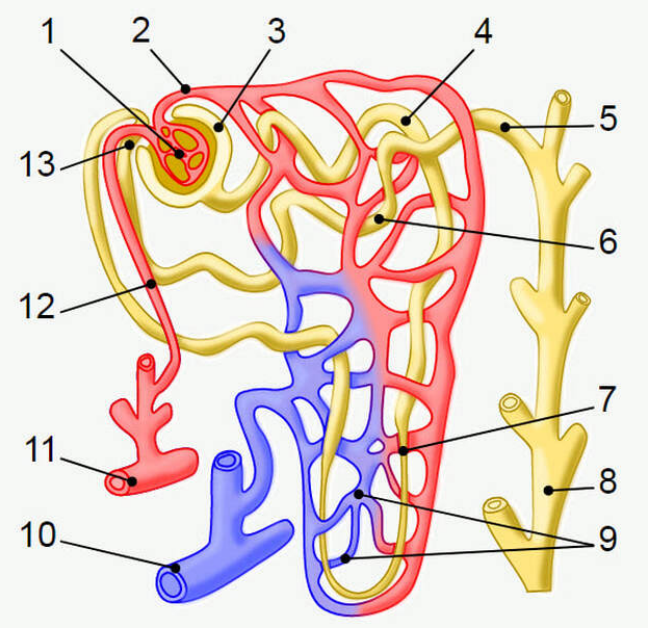


Figure 2.6.6.13.2.1. Components of the urinary system
(Image Source: Adapted from iStock Photos, ©)

Three processes are involved in urinary excretion:

1. Filtration
2. Secretion

3. Reabsorption



Legend:
1. Glomerulus
2. Efferent arteriole
3. Bowman's capsule
4. Proximal convoluted tubule
5. Cortical collecting duct
6. Distal convoluted tubule
7. Loop of Henle
8. Papillary duct
9. Peritubular capillaries
10. Arcuate vein
11. Arcuate artery
12. Afferent arteriole
13. Juxtaglomerular apparatus

Figure 2.6.6.13.2.2. Nephron of the kidney
(Image Source: Adapted from Wikimedia Commons, obtained by Public Domain, Creative Commons CC0 1.0 Universal Public Domain Dedication.
View original image.)

Filtration

Filtration takes place in the glomerulus, which is the vascular beginning of the nephron. Approximately one-fourth of the blood flow from cardiac output circulates through the kidney, the greatest rate of blood flow for any organ. A considerable amount of the blood plasma filters through the glomerulus into the nephron tubule. This results from the large amount of blood flow through the glomerulus, the large pores (40 Angstrom [Å]) in the glomerular capillaries, and the hydrostatic pressure of the blood. Small molecules, including water, readily pass through the sieve-like filter into the nephron tubule. Both lipid soluble and polar substances will pass through the glomerulus into the tubule filtrate. The amount of filtrate is very large, about 45 gallons per day in an adult human. About 99% of the water-like filtrate, small molecules, and lipid-soluble substances, are reabsorbed downstream in the nephron tubule. This means that the amount of urine eliminated is only about one percent of the amount of fluid filtrated through the glomeruli into the renal tubules.

Molecules with molecular weights greater than 60,000 (which include large protein molecules and blood cells) cannot pass through the capillary pores and remain in the blood. If urine contains albumin or blood cells, it indicates that the glomeruli have been damaged. Binding to plasma proteins will influence urinary excretion. Polar substances usually do not bind with the plasma proteins and thus can be filtered out of the blood into the tubule filtrate. In contrast, substances extensively bound to plasma proteins remain in the blood.

Secretion

Secretion, which occurs in the proximal tubule section of the nephron, is responsible for the transport of certain molecules out of the blood and into the urine. Secreted substances include potassium ions, hydrogen ions, and some xenobiotics. Secretion occurs by active transport mechanisms that are capable of differentiating among compounds based on polarity. Two systems exist, one that transports **weak acids** (such as many conjugated drugs and penicillins) and the other that transports **basic substances** (such as histamine and choline).

Reabsorption

Reabsorption takes place mainly in the proximal convoluted tubule of the nephron. Nearly all of the water, glucose, potassium, and amino acids lost during glomerular filtration reenter the blood from the renal tubules. Reabsorption occurs primarily by passive transfer based on a concentration gradient, moving from a high concentration in the proximal tubule to the lower concentration in the capillaries surrounding the tubule (Figures 4-6).

A factor that greatly affects reabsorption and urinary excretion is the pH of the urine. This is especially the case with weak electrolytes. If the urine is alkaline, weak acids are more ionized and excretion is increased. Weak acids (such as glucuronide and sulfate conjugates) are less ionized if the urine is acidic and undergo reabsorption and renal excretion is reduced. Since the urinary pH varies in humans, the urinary excretion rates of weak electrolytes also vary.

- Examples are phenobarbital (an acidic drug) which is ionized in alkaline urine and amphetamine (a basic drug) which is ionized in acidic urine. Treatment of barbiturate poisoning (such as an overdose of phenobarbital) may include changing the pH of the urine to facilitate excretion.
- Diet may have an influence on urinary pH and thus the elimination of some toxicants. For example, a high-protein diet results in acidic urine.

The physical properties (primarily molecular size) and polarity of a substance in the urinary filtrate greatly affect its ultimate elimination by the kidney. Small toxicants (both polar and lipid-soluble) are filtered with ease by the glomerulus. In some cases, large molecules (including some that are protein-bound) may be secreted (by passive transfer) from the blood across capillary endothelial cells and nephron tubule membranes to enter the urine. The major difference in ultimate fate is governed by a substance's polarity. Those substances that are ionized remain in the urine and leave the body. Lipid-soluble toxicants can be reabsorbed and re-enter the blood circulation, which lengthens their half-life in the body and potential for toxicity.

Kidneys, which have been damaged by toxins, infectious diseases, or because of age, have diminished ability to eliminate toxicants thus making those individuals more susceptible to toxins that enter the body. The presence of albumin in the urine indicates that the glomerulus filtering system is damaged, letting large molecules pass through. The presence of glucose in the urine is an indication that tubular reabsorption has been impaired.

Knowledge Check

1) The reason that much of the blood plasma filters into the renal tubule is due to:

a) The large amount of blood, under relatively high pressure, that flows through kidney glomerulae whose capillaries have large pores
b) Its high lipid content
c) The high binding content of plasma

Answer

The large amount of blood, under relatively high pressure, that flows through kidney glomerulae whose capillaries have large pores - **This is the correct answer.**
A considerable amount of the blood plasma filters through the glomerulus into the nephron tubule. This results from the large amount of blood flow through the glomerulus, the large pores (40 Angstrom [Å]) in the glomerular capillaries, and the hydrostatic pressure of the blood.

2) In which area of the nephron does active secretion take place?

a) The collecting duct of the nephron
b) The proximal tubule of the nephron
c) The glomerulus of the nephron

Answer

The proximal tubule of the nephron - **This is the correct answer.**
Secretion occurs in the proximal tubule section of the nephron and is responsible for the transport of certain molecules out of the blood and into the urine.

3) Most of the material filtered through the glomerulus is reabsorbed in the proximal convoluted tubule of the nephron. The primary property of a xenobiotic that determines whether it will be reabsorbed is:

a) Protein binding
b) Molecular size
c) Polarity

Answer

Polarity - **This is the correct answer.**
The ultimate fate of a substance filtered into the renal tubule is governed by its polarity. Those substances that are ionized remain in the urine and leave the body.

This page titled 2.6.6.13.2: Urinary Excretion is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **13.2: Urinary Excretion** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.13.3: Fecal Excretion

Fecal Excretion

Elimination of toxicants in the feces occurs from two processes:

1. Excretion in bile, which then enters the intestine ("biliary excretion").
2. Direct excretion into the lumen of the gastrointestinal tract ("intestinal excretion").

Biliary Excretion

The **biliary route** is an important mechanism for fecal excretion of xenobiotics and is even more important for the excretion of their metabolites. This route generally involves active secretion rather than passive diffusion. Specific transport systems appear to exist for certain types of substances, for example, organic bases, organic acids, and neutral substances. Some heavy metals are excreted in the bile, for example, arsenic, lead, and mercury. However, the most likely substances to be excreted via the bile are comparatively large, ionized molecules, such as those having a large molecular weight (conjugates greater than 300).

Once a substance has been excreted by the liver into the bile, and then into the intestinal tract, it can be eliminated from the body in the feces, or it may be reabsorbed. Since most of the substances excreted in the bile are water soluble, they are not likely to be reabsorbed as such. However, enzymes in the intestinal flora are capable of hydrolyzing some glucuronide and sulfate conjugates, which can release the less polar compounds that may then be reabsorbed. This process of excretion into the intestinal tract via the bile and reabsorption and return to the liver by the portal circulation is known as the **enterohepatic circulation** (Figure 1).

Enterohepatic circulation prolongs the life of the xenobiotic in the body. In some cases, the metabolite is more toxic than the excreted conjugate. Continuous enterohepatic recycling can occur and lead to very long half-lives of some substances. For this reason, drugs may be given orally to bind substances excreted in the bile.

- For example, a resin can be taken orally to bind with dimethylmercury, which had been secreted in the bile. The binding of the resin to dimethylmercury prevents its reabsorption and further toxicity.

Changes in the production and flow of bile into the liver affect the efficiency of biliary excretion.

- Liver disease usually causes a decrease in bile flow.
- Some drugs such as phenobarbital can produce an increase in bile flow rate. Administration of phenobarbital has been shown to enhance the excretion of methylmercury by this mechanism.

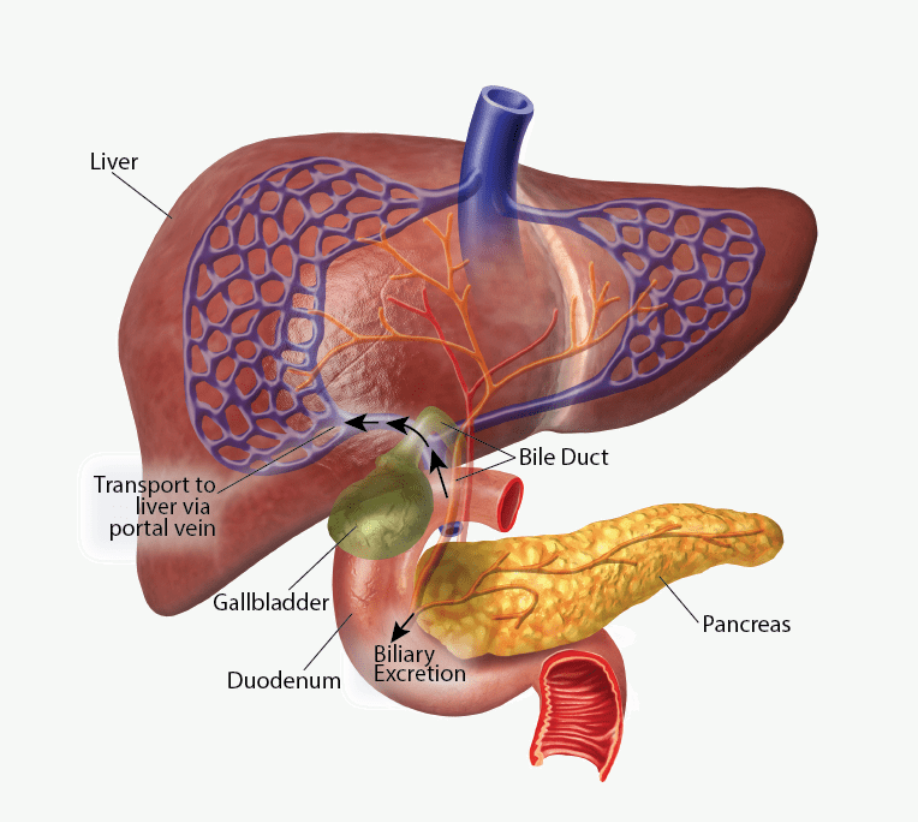


Figure 2.6.6.13.3.1. Biliary excretion and enterohepatic circulation
(Image Source: Adapted from iStock Photos, ©)

Intestinal Excretion

Another way that xenobiotics can be eliminated via the feces is by direct **intestinal excretion**. While this is not a major route of elimination, a large number of substances can be excreted into the intestinal tract and eliminated via feces. Some substances, especially those that are poorly ionized in plasma (such as weak bases), may passively diffuse through the walls of the capillaries, through the intestinal submucosa, and into the intestinal lumen to be eliminated in feces.

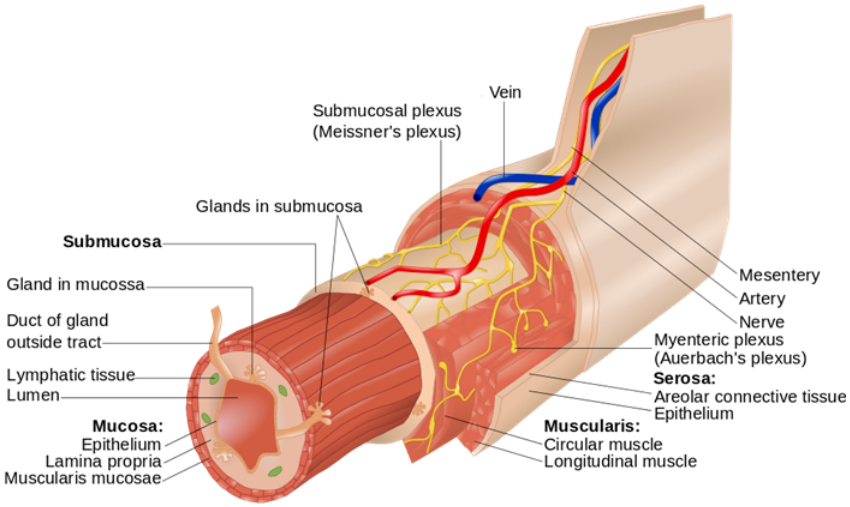


Figure 2.6.6.13.3.2. Layers of the Alimentary Canal
(Image Source: Wikimedia Commons, obtained under Creative Commons Attribution-Share Alike 3.0 Unported license. Author: Goran tek-en.
View original image.)

Intestinal excretion is a relatively slow process and therefore, it is an important elimination route only for those xenobiotics that have slow biotransformation, or slow urinary or biliary excretion. Increasing the lipid content of the intestinal tract can enhance intestinal excretion of some lipophilic substances. For this reason, mineral oil (liquid paraffin, derived from petroleum) is sometimes added to the diet to help eliminate toxic substances, which are known to be excreted directly into the intestinal tract.

Knowledge Check

1) Substances excreted in the bile are primarily:

a) Small, lipid soluble molecules
b) Comparatively large, ionized molecules
c) Large, lipid soluble molecules

Answer

Comparatively large, ionized molecules - **This is the correct answer.**
The most likely substances to be excreted via the bile are comparatively large, ionized molecules, such as large molecular weight (greater than 300) conjugates.

2) Many substances excreted in bile undergo enterohepatic circulation, which involves:

a) Excretion of substances into the circulating system rather than into the intestine
b) Excretion into the intestinal tract and reabsorption and return to the liver by the portal circulation
c) The recycling of xenobiotics between the liver and gall bladder

Answer

Excretion into the intestinal tract and reabsorption and return to the liver by the portal circulation - **This is the correct answer.**
The process of excretion into the intestinal tract via the bile and reabsorption and return to the liver by the portal circulation is known as the enterohepatic circulation. The effect of this enterohepatic circulation is to prolong the life of the xenobiotic in the body.

This page titled 2.6.6.13.3: Fecal Excretion is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **13.3: Fecal Excretion** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.13.4: Exhaled Air

Exhaled Air

The lungs are an important route of excretion for xenobiotics (and metabolites) that exist in a gaseous phase in the blood.

Passive Diffusion

Blood gases are excreted by passive diffusion from the blood into the alveolus, following a concentration gradient. This type of excretion occurs when the concentration of the xenobiotic dissolved in capillary blood is greater than the concentration of the substance in the alveolar air. Gases with a low solubility in blood are more rapidly eliminated than those gases with a high solubility. Volatile liquids dissolved in the blood are also readily excreted via the expired air.

For example, breathalyzer devices can measure blood alcohol concentration because as alcohol in the blood moves across the alveoli the alcohol in the blood evaporates and is exhaled. The concentration of alcohol in the exhaled air relates to the level of alcohol in the blood.

Impact of Vapor Pressure

The amount of a liquid excreted by the lungs is proportional to its vapor pressure. Exhalation is an exception to most other routes of excretion in that it can be a very efficient route of excretion for lipid soluble substances. This is due to the very close proximity of capillary and alveolar membranes, which are thin and allow for the normal gaseous exchange that occurs in breathing.

Knowledge Check

1) Xenobiotics are eliminated in exhaled air by:

a) Passive diffusion
b) Active transport
c) Facilitated transport

Answer

Passive diffusion - **This is the correct answer.**
Blood gases are excreted by passive diffusion from the blood into the alveolus, following a concentration gradient. This occurs when the concentration of the xenobiotic dissolved in capillary blood is greater than the concentration of the substance in the alveolar air.

This page titled 2.6.6.13.4: Exhaled Air is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **13.4: Exhaled Air** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.13.5: Other Routes

Other Routes of Excretion

Several minor routes of excretion occur including mother's milk, sweat, saliva, tears, and semen.

Excretion into Breast Milk

Excretion into milk can be important since toxicants can be passed with milk to the nursing offspring. In addition, toxic substances can pass from cow's milk to people. Toxic substances are excreted into milk by simple diffusion. Both basic substances and lipid soluble compounds can be **excreted into milk (**The National Library of Medicine's **LactMed** is a resource for information on drugs, dietary supplements, and herbs that pass into breast milk.).

Basic substances can be concentrated in milk since milk is more acidic (pH approximately 6.5) than blood plasma. Since milk contains 3–4% lipids, lipid soluble xenobiotics can diffuse along with fats from plasma into the mammary gland and thus can be present in mother's milk. Substances such as lead, mercury, Bisphenol A (BPA), and phthalates that are chemically similar to calcium can also be excreted into milk along with calcium.

Did you know?

Volatile organic compounds (VOCs) found in indoor air can also be found in breast milk.

Examples include MTBE (methyl tert-butyl ether), chloroform, benzene, and toluene. For benzene, toluene, and MTBE, the levels in breast milk followed the indoor air concentrations. However, the infant average daily dose by inhalation exceeded ingestion rates by 25-to-135 fold. Thus, the amount of VOC exposure from indoor air in nonsmoking households is much greater than the VOC exposure from breast milk. Strategies to lessen infant VOC exposure should focus on improving indoor air quality.

Excretion into All Other Body Secretions or Tissues

Excretion of xenobiotics in **all other body secretions or tissues** (including the saliva, sweat, tears, hair, and skin) are of only minor importance. Under conditions of great sweat production, excretion in sweat may reach a significant degree. Some metals, including cadmium, copper, iron, lead, nickel, and zinc, may be eliminated in sweat to some extent. Xenobiotics that passively diffuse into saliva may be swallowed and absorbed by the gastrointestinal system. The excretion of some substances into saliva is responsible for the unpleasant taste that sometimes occurs with time after exposure to a substance.

Knowledge Check

1) The following are minor routes of excretion:

a) Sweat and saliva
b) Urinary excretion, fecal excretion, and exhaled air

Answer

Sweat and saliva - **This is the correct answer.**
Several minor routes of excretion exist, primarily via mother's milk, sweat, saliva, tears, and semen. The main routes of excretion are via urine, feces, and exhaled air.

This page titled 2.6.6.13.5: Other Routes is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **13.5: Other Routes** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.14: Cellular Toxicology

Learning Objectives

After completing this lesson, you will be able to:

- Explain cellular adaptation.
- Identify four possible endpoints to toxic damage to cells and tissues.
- Define commonly used cancer terms.
- Describe the phases of and genetic activity associated with carcinogenesis.
- Identify mechanisms and potential outcomes of neurotoxicity.

In this section...

Topics include:

2.6.6.14.1: Adaptation

2.6.6.14.2: Cell Damage and Tissue Repair

2.6.6.14.3: Cancer

2.6.6.14.4: Neurotoxicity

Section 14: Key Points

What We've Covered

This section made the following main points:

- To maintain homeostasis, cells and tissues undergo:
  - Physiological adaptation, which is beneficial in nature — for example, increased skeletal muscle cells in athletes.
  - Pathological adaptation, which is detrimental — for example, cellular changes in people who smoke cigarettes.
- Specific types of adaptation include:
  - Atrophy — a decrease in the size of cells.
  - Hypertrophy — an increase in the size of individual cells.
  - Hyperplasia — an increase in the number of cells in a tissue.
  - Metaplasia — the conversion from one type of mature cell to another type.
  - Dysplasia — abnormal cell changes or deranged cell growth.
  - Anaplasia — cells that are undifferentiated.
  - Neoplasia — new growth of tissue.
- Most toxic effects, especially due to xenobiotics, are due to specific biochemical interactions without causing recognizable damage to a cell or its organelles. Cellular or biochemical toxicity leads to:
  - The tissue being completely repaired and returned to normal.

- The tissue being incompletely repaired but capable of functioning with reduced capacity.
- Death of the organism or complete loss of a tissue or organ.
- Neoplasm or cancers.

- Tumors are either:
  - Benign — similar to the cell of origin, slow-growing, and usually without systemic effects.
  - Malignant — dissimilar from the cell of origin, rapid-growing, and commonly with systemic effects and life-threatening. Most malignant tumors are either:
    - Carcinomas — arising in epithelium, the most common form of cancer, usually spread in the lymphatic system.
    - Sarcomas — arising in connective or muscle tissue, usually spread by the blood stream.
- Carcinogenesis is a multi-step, multi-factorial genetic disease consisting of at least three main phases:
  1. Initiation — irreversible alteration of the DNA (mutation) of a normal cell.
  2. Promotion/Conversion — promoters enhance further development of the initiated cells, often influencing further expression of the mutated DNA such that the initiated cell proliferates and progresses further.
  3. Progression — development of the initiated cell into a biologically malignant cell population, often with metastasis to other areas of the body.
- Regulatory genes control the activity of structural genes and direct the proliferation process of the cell. Regulatory genes that play roles in carcinogenesis include:
  - Proto-oncogenes — normal cellular genes that encode and instruct the production of regulatory proteins and growth factors within a cell or its membrane.
  - Oncogenes — altered or misdirected proto-oncogenes with the ability to direct the production of proteins within the cell that change or transform the normal cell into a neoplastic cell.
  - Tumor suppressor genes (anti-oncogenes) — present in normal cells and counteract and change the proto-oncogenes and altered proteins, preventing a cell with damaged DNA from proliferating and evolving into an uncontrolled growth.
    - The p53 gene normally halts cell division, stimulates repair enzymes, and if necessary, commands the mutated cell to self-destruct
    - p53 is the most frequently altered in human tumors and is incapable of its defense mechanisms
- Toxic damage to the nervous system is divided into three categories:
  1. Damage to sensory receptors and sensory neurons impacting the sensory functions.
  2. Damage to motor neurons causing muscular weakness and paralysis.
  3. Interneuronal damage causing learning deficiencies, memory loss, incoordination, and emotional conditions.

This page titled 2.6.6.14: Cellular Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.6.14.1: Adaptation

This section discusses cellular effects yet cell and chemical effects cannot be conveniently separated because cells are constructed of a variety of chemicals of diverse types. Specific intracellular chemical changes may occur as changes in the cell and may affect either its appearance or function. The actual mechanisms leading to cell damage are usually biochemical in nature.

Adaptation Explained

To maintain homeostasis, cells and tissues:

- "Cope" with new demands placed on them by constantly adapting to changes in the tissue environment.
- Are usually capable of an amazing degree of cellular adaptability.
- Adapt in a way that may be beneficial in nature (physiological) or detrimental (pathological).

Examples of **physiological adaptation** are:

- An increase in skeletal muscle cells in athletes due to exercise and increased metabolic demand.
- The increase in number and size of epithelial cells in breasts of women resulting from endocrine stimulation during pregnancy.

When these cells or tissues are damaged, the body attempts to adapt and repair or limit the harmful effects. Often the adaptive changes result in cells or organs that cannot function normally. This imperfect adaptation is a pathological change.
Examples of **pathological adaptations** are:

- Cellular changes in people who smoke cigarettes: The ciliated columnar epithelium changes to non-ciliated squamous epithelium in the trachea and bronchi of cigarette smokers. The replacement of squamous epithelium can better withstand the irritation of the cigarette smoke. However, the loss of cilia and mucous secretions of columnar epithelium diminish the tracheobronchial defense mechanisms.
- Replacement of normal liver cells by fibrotic cells in chronic alcoholics (known as cirrhosis of the liver): A severely cirrhotic liver is incapable of normal metabolism, maintenance of nutrition, and detoxification of xenobiotics.

If the change is minor, cellular adaptation may result and the cells return to normal. When damage is very severe, the result may be cell death or permanent functional incapacitation.
Cellular adaptation to toxic agents includes three basic types:

1. Increase in cell activity.
2. Decrease in cell activity.
3. Alteration in cell morphology (structure and appearance) or cell function.

Specific Types of Cellular Adaptations

Atrophy

Atrophy is a **decrease in the size of cells**. If a sufficient number of cells are involved, the tissue or organ may also decrease in size. When cells atrophy, they have:

- Reduced oxygen needs.
- Reduced protein synthesis.
- Decreased number and size of the organelles.

The most common causes of atrophy are reduced use of the cells, lack of hormonal or nerve stimulation, decrease in nutrition, reduced blood flow to the tissue, and natural aging.

- An example of atrophy is the decrease in the size of muscles and muscle cells in persons whose legs are paralyzed, in a cast, or infrequently used as when a patient is on bedrest.

Hypertrophy

Hypertrophy is an **increase in the size of individual cells**. This frequently results in an increase in the size of a tissue or organ. When cells hypertrophy, components of the cell increase in numbers with increased functional capacity to meeting increased cell needs. Hypertrophy generally occurs in situations where the organ or tissue cannot adapt to an increased demand by formation of more cells. This is commonly seen in cardiac and skeletal muscle cells, which do not divide to form more cells. Common causes for hypertrophy are increased work or stress placed on an organ or hormonal stimulation.

- An example of hypertrophy is the compensatory increase in the size of cells in one kidney after the other kidney has been removed or is in a diseased state.

Hyperplasia

Hyperplasia is an **increase in the number of cells in a tissue**. This generally results in an enlargement of tissue mass and organ size. It occurs *only* in tissues capable of mitosis such as the epithelium of skin, intestine, and glands. Some cells do not divide and thus cannot undergo hyperplasia, for example, nerve and muscle cells. Hyperplasia is often a compensatory measure to meet an increase in body demands. Hyperplasia is a frequent response to toxic agents and damage to tissues such as wounds or trauma. In wound healing, hyperplasia of connective tissue (for example, fibroblasts and blood vessels) contributes to the wound repair. In many cases, when the toxic stress is removed, the tissue returns to normal. Hyperplasia may result from hormonal stimulation, for example, breast and uterine enlargement due to increased estrogen production during pregnancy.

Metaplasia

Metaplasia is the **conversion** from one type of mature cell to another type of mature cell. It is a cellular replacement process. A metaplastic response often occurs with chronic irritation and inflammation. This results in a tissue more resistant to the external stress since the replacement cells are capable of survival under circumstances in which the original cell type could not survive. However, the cellular changes usually result in a loss of function, which was performed by the original cells that were lost and replaced.
Examples of metaplasia are:

- The common condition in which a person suffers from chronic reflux of acid from the stomach into the esophagus (Gastroesophageal Reflux Disease). The normal esophageal cells (squamous epithelium) are sensitive to the refluxed acid and die. They are replaced with the columnar cells of the stomach that are resistant to the stomach's acidity. This pathological condition is known as "Barrett's Esophagus."
- The change in the cells of the trachea and bronchi of chronic cigarette smokers from ciliated columnar epithelium to non-ciliated stratified squamous epithelium. The sites of metaplasia frequently are also sites for neoplastic transformations. The replacement cells lack the defense mechanism performed by the cilia in moving particles up and out of the trachea.
- With cirrhosis of the liver, which is a common condition of chronic alcoholics, the normal functional hepatic cells are replaced by nonfunctional fibrous tissue.

Dysplasia

Dysplasia is a condition of **abnormal cell changes or deranged cell growth** in which the cells are structurally changed in size, shape, and appearance from the original cell type. Cellular organelles also become abnormal. A common feature of dysplastic cells is that the nuclei are larger than normal and the dysplastic cells have a mitotic rate higher than the predecessor normal cells. Causes of dysplasia include chronic irritation and infection. In many cases, the dysplasia can be reversed if the stress is removed and normal cells return. In other cases, dysplasia may be permanent or represent a precancerous change.

- An example of dysplasia is the atypical cervical cells that precede cervical cancer. Routine examination of cervical cells is a routine screening test for dysplasia and possible early stage cervical cancer (Papanicolaou test).
- Cancer occurs at the site of Barrett's syndrome and in the bronchi of chronic smokers (bronchogenic squamous cell carcinoma).

Anaplasia

Anaplasia refers to **cells that are undifferentiated**. They have irregular nuclei and cell structure with numerous mitotic figures. Anaplasia is frequently associated with malignancies and serves as one criterion for grading the aggressiveness of a cancer. For example, an anaplastic carcinoma is one in which the cell appearance has changed from the highly differentiated cell of origin to a cell type lacking the normal characteristics of the original cell. In general, anaplastic cells have lost the normal cellular controls, which regulate division and differentiation.

Neoplasia

Neoplasia is a **new growth of tissue** and is commonly referred to as a tumor. There are two types of neoplasia: benign and malignant. Malignant neoplasia are cancers. Since cancer is such an important and complex medical problem, a **separate section** is devoted to cancer.

Interactions

Interactions between two or more toxic agents can produce damage by chemical-chemical interactions, chemical-receptor interactions, or by modification, by a first agent, of the cell and tissue response to a second agent. Interactions may occur by simultaneous exposure and if exposure to the two agents is separated in time.

Chemical-chemical interactions have been mostly studied in the toxicology of air pollutants, where it was shown that the untoward effect of certain oxidants may be enhanced in the presence of other aerosols.

Interactions at the *receptor site* have been found in isolated perfused lung experiments. Oxygen tolerance may be an example, when pre-exposure to one concentration of oxygen mitigates later exposure to 100% oxygen by modifying cellular and enzymatic composition of the lung.

Damage of the alveolar zone by the antioxidant butylated hydroxytoluene (BHT) in mice can be greatly enhanced by *subsequent exposure* to oxygen concentration which, otherwise, would have little if any demonstrable effect.

The *synergistic interaction* between BHT and oxygen in mice results in interstitial pulmonary fibrosis. Acute or chronic lung disease may then be caused not only by one agent, but also in many instances by the interaction of several agents.

Knowledge Check

1) An increase in skeletal muscle cells in athletes due to exercise and increased metabolic demand is an example of:

a) Pathological adaptation
b) Physiological adaptation

Answer

Physiological adaptation - **This is the correct answer.**
The increase in skeletal muscle cells in athletes due to exercise and increased metabolic demand is an example of physiological adaptation since the increased muscle is beneficial rather than harmful.

2) A cellular response in which there is an increase in the number of cells in a tissue is known as:

a) Atrophy
b) Hypertrophy

c) Hyperplasia
d) Metaplasia

Answer

Hyperplasia - **This is the correct answer.**
Hyperplasia is an increase in the number of cells in a tissue.

3) A condition of abnormal cell changes or deranged cell growth in which the cells are structurally changed in size, shape, and appearance from the original cell type is known as:

a) Dysplasia
b) Anaplasia
c) Neoplasia

Answer

Dysplasia - **This is the correct answer.**
Dysplasia is a condition of abnormal cell changes or deranged cell growth in which the cells are structurally changed in size, shape, and appearance from the original cell type.

This page titled 2.6.6.14.1: Adaptation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **14.1: Adaptation** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.14.2: Cell Damage and Tissue Repair

Cell Damage and Tissue Repair

Toxic damage to cells can cause individual cell death and if sufficient cells are lost, the result can be tissue or organ failure, ultimately leading to death of the organism. It is nearly impossible to separate a discussion of cellular toxicity and biochemical toxicity. Most observable cellular changes and cell death are due to specific biochemical changes within the cell or in the surrounding tissue. However, there are a few situations where a toxic chemical or physical agent can cause cell damage without actually affecting a specific chemical in the cell or its membrane. Physical agents such as heat and radiation may damage a cell by coagulating their contents (similar to cooking). In this case, there are no specific chemical interactions. Impaired nutrient supply (such as glucose and oxygen) may deprive the cell of essential materials needed for survival.

Toxic Effects

The majority of toxic effects, especially due to xenobiotics, are due to specific biochemical interactions without causing recognizable damage to a cell or its organelles.
Examples of these toxic effects include:

- Interference with a chemical that transmits a message across a neural synapse such as the inhibition of the enzyme acetylcholinesterase by organophosphate pesticides.
- When one toxic chemical inhibits or replaces another essential chemical such as the replacement of oxygen on the hemoglobin molecule with carbon monoxide.

The human body is extremely complex. In addition to over 200 different cell types and about as many types of tissues, there are literally thousands of different biochemicals, which may act alone or in concert to keep the body functions operating correctly. To illustrate the cell's structures and functions and the chemical toxicity of all tissues and organs would be impossible in this brief tutorial. This section presents only a general overview of toxic effects along with some specific types of toxicity that include cancer and neurotoxicity.

Capacity for Repair

Some tissues have a great capacity for repair, such as most epithelial tissues. Others have limited or no capacity to regenerate and repair, such as nervous tissue. Most organs have a functional reserve capacity so that they can continue to perform their body function although perhaps in somewhat diminished ability. For example:

- Half of a person's liver can be damaged, and the body can regenerate sufficient new liver or repair the damaged section by fibrous replacement to maintain most of the capacity of the original liver.
- The hypertrophy of one kidney to assume the capacity lost when the other kidney has been lost or surgically removed.

Toxic Damage to Cells and Tissues

Toxic damage to cells and tissues can be transient and non-lethal or, in severe situations, the damage may cause death of the cells or tissues. The following diagram illustrates the various effects that can occur with damage to cells. There are four main final endpoints to the cellular or biochemical toxicity:

1. The tissue may be completely repaired and return to normal.
2. The tissue may be incompletely repaired but is capable of sustaining its function with reduced capacity.
3. Death of the organism or the complete loss of a tissue or organ. In some instances, the organism can continue to live with the aid of medical treatment, for example, replacement of insulin or by organ transplantations.
4. Neoplasm or cancers may result, many of which will result in death of the organism and some of which may be cured by medical treatment.

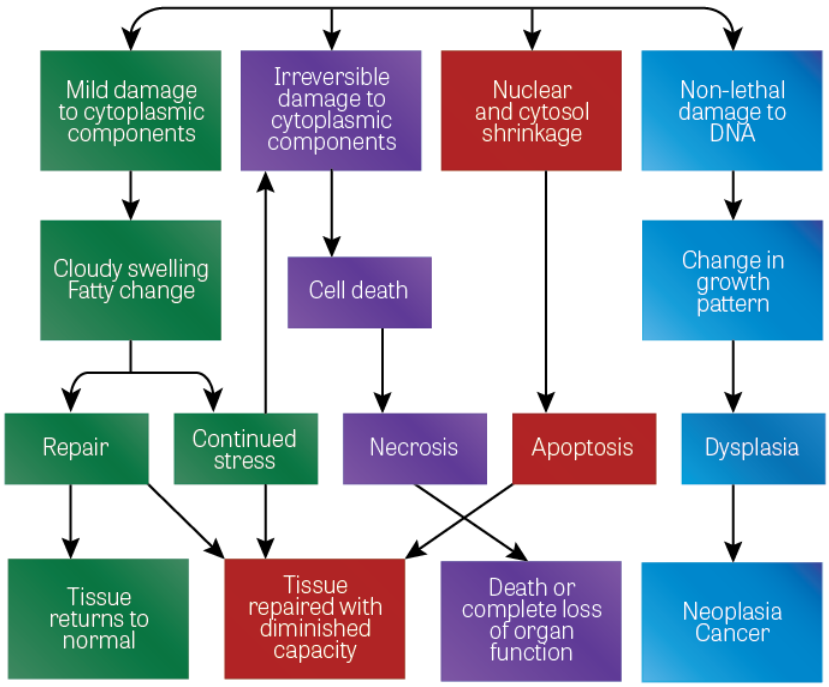


Figure 2.6.6.14.2.1. Toxic damage to cells
(Image Source: NLM)

Reversible Cell Damage

The response of cells to toxic injury may be transient and reversible once the stress has been removed or the compensatory cellular changes are made. In some cases, the full capability of the damaged cells returns. In other cases, a degree of permanent injury remains with a diminished cellular or tissue capacity. In addition to the adaptive cell changes discussed previously, two commonly encountered specific cell changes are associated with toxic exposures, cellular swelling, and fatty change.

Cellular swelling, which is associated with hypertrophy, is due to cellular hypoxia, which damages the sodium-potassium membrane pump. This in turn changes the intracellular electrolyte balance with an influx of fluids into the cell, causing it to swell. Cell swelling is reversible when the cause is eliminated.

Fatty change is more serious and occurs with severe cellular injury. In this situation, the cell has become damaged and is unable to adequately metabolize fat. The result is that small vacuoles of fat accumulate and become dispersed within the cytoplasm. While fatty change can occur in several organs, it is usually observed in the liver. This is because most fat is synthesized and metabolized in liver cells. Fatty change can be reversed but it is a much slower process than the reversal of cellular swelling.

Lethal Injury (Cell Death)

In many situations, the damage to a cell may be so severe that the cell cannot survive. Cell death occurs mainly by two methods: necrosis and apoptosis.

Necrosis is a progressive failure of essential metabolic and structural cell components usually in the cytoplasm. Necrosis generally involves a group of contiguous cells or occurs at the tissue level. Such progressive deterioration in structure and function rapidly leads to cell death or "necrotic cells." Necrosis begins as a reduced production of cellular proteins, changes in electrolyte gradient, or loss of membrane integrity (especially increased membrane permeability). Cytoplasmic organelles (such as mitochondria and

endoplasmic reticulum) swell while others (especially ribosomes) disappear. This early phase progresses to fluid accumulation in the cells making them pale-staining or showing vacuoles, which pathologists call "cloudy swelling" or "hydropic degeneration." In some cells, they no longer can metabolize fatty acids so that lipids accumulate in the cytoplasmic vacuoles, referred to as "fatty accumulation" or "fatty degeneration." In the final stages of "cell dying," the nucleus becomes shrunken (pyknosis) or fragmented (karyorrhexis).

Apoptosis or "programmed cell death" is a process of self-destruction of the cell nucleus. Apoptosis is an individual or single cell death in that dying cells are not contiguous but are scattered throughout a tissue. Apoptosis is a normal process in cell turnover in that cells have a finite lifespan and spontaneously die. During embryonic development, certain cells are programmed to die and are not replaced, such as the cells between each developing finger. If the programmed cells do not die, the fetus ends up with incomplete or fingers joined together in a web fashion.

In apoptosis, the cells shrink from a decrease of cytosol and the nucleus. The organelles (other than the nucleus) appear normal in apoptosis. The cell disintegrates into fragments referred to as "apoptotic bodies." These apoptotic bodies and the organelles are phagocytized by adjacent cells and local macrophages without initiation of an inflammatory response as is seen in necrosis. The cells undergo apoptosis and just appear to "fade away." Some toxicants induce apoptosis or, in other cases, they inhibit normal physiological apoptosis.

Following necrosis, the tissue attempts to regenerate with the same type of cells that have died. When the injury is minimal, the tissue may effectively replace the damaged or lost cells. In severely damaged tissues or long-term chronic situations, the ability of the tissue to regenerate the same cell types and tissue structure may be exceeded, so that a different and imperfect repair occurs.

- An example of this is with chronic alcoholic damage to liver tissue in which the body can no longer replace hepatocytes with hepatocytes but rather connective tissue replacement occurs. Fibrocytes with collagen replace the hepatocytes and normal liver structure with scar tissue. The fibrotic scar tissue shores up the damage but it cannot replace the function of the lost hepatic tissue. With constant fibrotic change, the liver function is continually diminished so that eventually the liver can no longer maintain homeostasis. This fibrotic replacement of the liver is known as cirrhosis (Figure 2.6.6.14.2.2). The normal dark-red, glistening smooth appearance of the liver has been replaced with light, irregular fibrous scar tissue that permeates the entire liver.

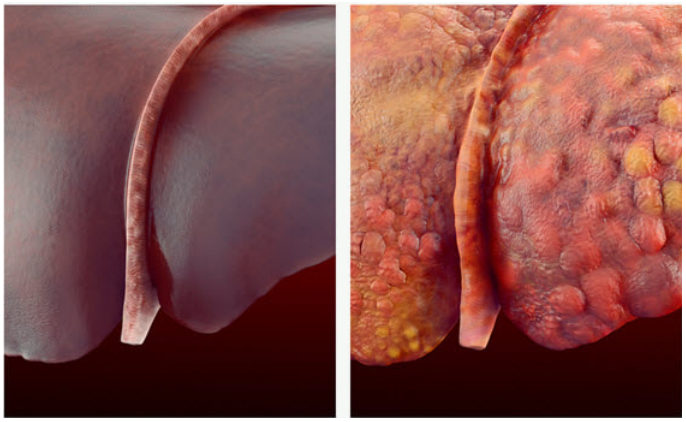


Figure 2.6.6.14.2.2. A healthy liver (left) and a liver with cirrhosis (right)
(Image Source: iStock Photos, ©)

We have so far discussed primarily changes to individual cells. However, a tissue and an organ consist of different types of cells that work together to achieve a particular function. As with a football team, when one member falters, the others rally to compensate. It is the same with a tissue. Damage to one cell type prompts reactions within the tissue to compensate for the injury. Within organs, there are two basic types of tissues: the parenchymal and stromal tissues. The **parenchymal tissues** contain the functional cells (for example, squamous dermal cells, liver hepatocytes, and pulmonary alveolar cells). The **stromal cells** are the supporting connective tissues (for example, blood vessels and elastic fibers).

Cell Repair

Repair of injured cells can be accomplished by either:

1. Regeneration of the parenchymal cells.
2. Repair and replacement by the stromal connective tissue.

The goal of the repair process is to fill the gap that results from the tissue damage and restore the structural continuity of the injured tissue. Normally a tissue attempts to regenerate the same cells that are damaged; however, in many cases, this cannot be achieved so that replacement with a stromal connective tissue is the best means for achieving the structural continuity.
The ability to **regenerate** varies greatly with the type of parenchymal cell. The regenerating cells come from the proliferation of nearby parenchymal cells, which serve to replace the lost cells. Based on regenerating ability, there are three types of cells:

1. **Labile cells** — cells that routinely divide and replace cells that have a limited lifespan (for example, skin epithelial cells, and hematopoietic stem cells).
2. **Stable cells** — cells that usually have a long lifespan with normally a low rate of division; they can rapidly divide upon demand.
3. **Permanent cells** — cells that never divide and do not have the ability for replication even when stressed or when some cells die.

Table 2.6.6.14.2.1 shows examples of cell types.

Cell Type	Examples
Labile cells	• Squamous epithelium of skin, mouth, vagina, and cervix • Columnar epithelium of intestinal tract • Transitional epithelium of urinary tract • Bone marrow cells
Stable cells	• Liver hepatocytes • Alveolar cells of lung • Epithelium of kidney tubules
Permanent cells	• Neurons • Skeletal and cardiac muscle

Table 1. Examples of three cell types of parenchymal cells

Cell Type	Examples
Labile cells	• Squamous epithelium of skin, mouth, vagina, and cervix • Columnar epithelium of intestinal tract • Transitional epithelium of urinary tract • Bone marrow cells
Stable cells	• Liver hepatocytes • Alveolar cells of lung • Epithelium of kidney tubules
Permanent cells	• Neurons • Skeletal and cardiac muscle

The labile cells have a great potential for regeneration by replication and repopulation with the same cell type so long as the supporting structure remains intact. Stable cells can also respond and regenerate but to a lesser degree and are quite dependent on the supporting stromal framework. When the stromal framework is damaged, the regenerated parenchymal cells may be irregularly

dispersed in the organ resulting in diminished organ function. The tissue response for the labile and stable cells is initially hyperplasia until the organ function becomes normal again. When permanent cells die they are not replaced in kind but instead connective tissue (usually fibrous tissue) moves in to occupy the damaged area. This is a form of metaplasia.
Examples of replacement by metaplasia are:

- **Cirrhosis of the liver** — liver cells (hepatocytes) are replaced by bands of fibrous tissue, which cannot carry out the metabolic functions of the liver.
- **Cardiac infarcts** — cardiac muscle cells do not regenerate and thus are replaced by fibrous connective tissue (scar). The scar cannot transmit electrical impulses or participate in contraction of the heart.
- **Pulmonary fibrosis** — damaged or dead epithelial cells lining the pulmonary alveoli are replaced by fibrous tissue. Gases cannot diffuse across the fibrous cells and thus gas exchange is drastically reduced in the lungs.

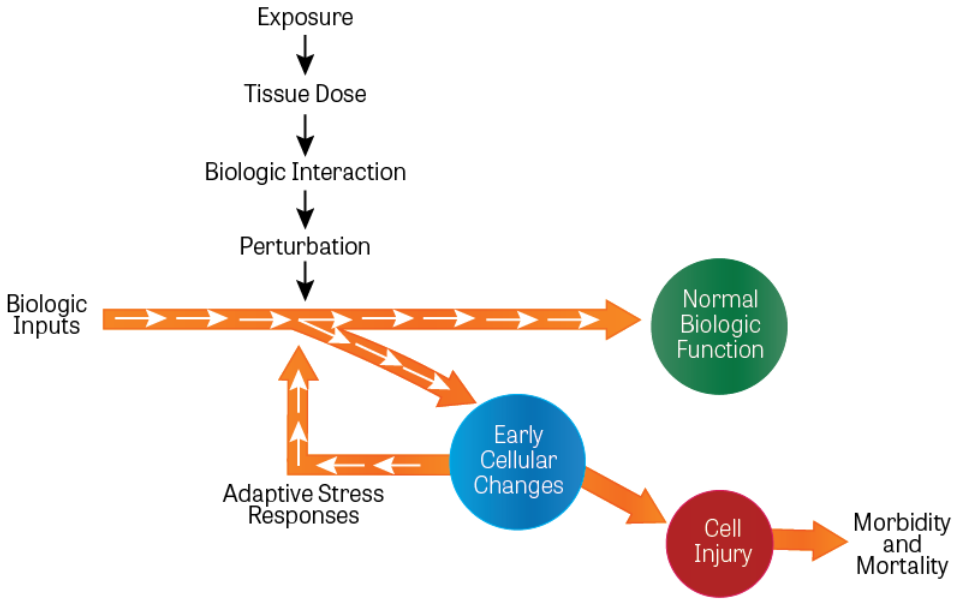


Figure 2.6.6.14.2.3. Activation of Toxicity Pathways
(Image Source: Adapted from Dr. Andrew Maier, adapted from National Research Council (NRC) 2007a.)

Knowledge Check

1) The process of self-destruction of the cell nucleus *(often referred to as "programmed cell death")* is known as:

a) Necrosis
b) Apoptosis
c) Cellular swelling
d) Fatty change

Answer

Apoptosis - **This is the correct answer.**
Apoptosis *(referred to as "programmed cell death")* is a process of self-destruction of the cell nucleus.

2) The category of cells that routinely divide and replace cells that have a limited lifespan is known as:

a) Labile cells
b) Stable cells
c) Permanent cells

Answer

Labile cells - **This is the correct answer.**
Labile cells are cells that routinely divide and replace cells that have a limited lifespan (e.g., skin epithelial cells and hematopoietic stem cells).

This page titled 2.6.6.14.2: Cell Damage and Tissue Repair is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **14.2: Cell Damage and Tissue Repair** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Cancer

Cancer has long been considered a cellular disease since cancers are composed of cells that grow without restraint in various areas of the body. Such growths of cancerous cells can replace normal cells or tissues causing severe malformations (such as with skin and bone cancers) and failure of internal organs which frequently leads to death. How do cells become cancerous? The development of cancer is an enormously complex process. For once a cell has started on the cancer path, it progresses through a series of steps, which continue long after the initial cause has disappeared.

Overview

There are about as many types of cancers as there are different types of cells in the body (over 100 types). Some cell types constantly divide and are replaced (such as skin and blood cells). Other types of cells rarely or never divide (such as bone cells and neurons). Sophisticated mechanisms exist in cells to control when, if, and how cells replicate. Cancer occurs when these mechanisms are lost and replication takes place in an uncontrolled and disorderly manner. It can arise when one cell or a small group of cells multiplies too many times because of damage to its DNA.

Recent research has begun to unravel the extremely complex pathogenesis of cancer. There is an intricate array of biochemical changes that take place within cells and between cells underlying the progression of cancer that transforms normal cells into cancerous cells. These biochemical changes lead a cell through a series of steps, changing it gradually from a normal to a cancer cell. The altered cell is no longer bound by the regulatory controls that govern the life and behavior of normal cells.

Cancer is not a single disease but a large group of diseases. The common aspect is that all cancers have the same basic property: they are composed of cells that no longer conform to the usual constraints on cell proliferation. In other words, they are uncontrolled growths of cells.

Terminology

The terminology associated with cancer can be confusing and may be used differently among the public and medical communities. Here are definitions of the most frequently used cancer terms:

- **Cancer** — a malignant tumor that has the ability to metastasize or invade into surrounding tissues.
- **Tumor** — a general term for an uncontrolled growth of cells that becomes progressively worse with time. Tumors may be benign or malignant.
- **Neoplasm** — same as a tumor.
- **Neoplasia** — the growth of new tissue with abnormal and unregulated cellular proliferation.
- **Benign Tumor** — a tumor that does not metastasize or invade surrounding tissue.
- **Malignant Tumor** — a tumor that has the ability to metastasize or invade into surrounding tissues (same as cancer).
- **Metastasis** — ability to establish secondary tumor growth at a new location away from the original site.
- **Carcinogenesis** — the production of a carcinoma (epithelial cancer). Sometimes carcinogenesis is used as a general term for production of any type of tumor.

How are Cancers Named?

While most tumors are generally named in accordance with an **internationally agreed-upon classification scheme**, there are exceptions. Tumors are generally named and classified based on:

- The cell or tissue of origin
- Whether benign or malignant

Most tumor names end with the suffix "oma" which indicates a swelling or tissue enlargement. [Note: some terms ending with -oma are not cancers; for example, a hematoma is merely a swelling consisting of blood].

In naming tumors, qualifiers may be added in addition to the tissue of origin and structural features. For example, a "poorly-differentiated bronchogenic squamous cell carcinoma" is a malignant tumor (carcinoma) of squamous cell type (original cell type), which arose in the bronchi of the lung (site where the cancer started), and in which the cancer cells are poorly differentiated, meaning they have lost much of the normal appearance of squamous cells.

There are several historical exceptions to the standard nomenclature system, often based on their early and accepted use in the literature.

Examples include:

- Some tumors are named after the person that first described the tumor, for example, Wilms tumor (kidney tumor) and Hodgkin lymphoma (a specific form of lymphoid cancer).
- A few cancers are named for their physical characteristics such as pheochromocytomas (dark-colored tumors of the adrenal gland).
- A few cancers are composed of mixtures of cells, for example, fibrosarcoma and carcinosarcoma.

Most malignant tumors fall into one of two categories: carcinomas or sarcomas. The major differences between carcinomas and sarcomas are listed in Table 2.6.6.14.3.1:

Carcinoma	Sarcoma
• Malignant tumor arising in epithelium • The most common form of cancer • Usually spread in lymphatic system	• Malignant tumor arising in connective or muscle tissue • Usually spread by blood stream • Frequently metastasizes to lung

Differences between Benign and Malignant Tumors

The biological and medical consequences of a tumor depend on whether it is benign or malignant.
Table 2.6.6.14.3.2 provides a comparison of the primary differences between benign and malignant tumors:

Characteristics	Benign Tumors	Malignant Tumors
Cell characteristics	Similar to cell of origin (well differentiated)	Dissimilar from cell of origin (poorly differentiated)
Growth characteristics	Tumor edges move outward in a smooth manner (encapsulated), grows by expansion, and compresses and displaces surrounding tissues Tumor cells stay attached to the clone or mass of cells and do not break away and start new growths elsewhere in the body	Tumor edges move outward in an irregular fashion (usually no capsule) and can infiltrate, invade, and destroy surrounding tissues Tumor cells can break away from the cloned mass, live independently, move to other areas of the body, and start new clones or growths
Rate of growth	Slow	Rapid
Degree of vascularity	Slight	Moderate-marked
Recurrence after surgical removal	Seldom	Frequently
Degree of necrosis and ulceration	Unusual	Common
Likelihood of causing systemic effects	Unusual unless the tumor is a secreting endocrine neoplasm	Common and usually life-threatening

Common Sites for Cancer

Cancer can occur in almost any tissue or organ. Some cells and tissues are more likely to become cancerous than others are, particularly those **cells that normally undergo proliferation** to replace cells that have been lost due to injury or cell death. Those cells that don't proliferate (for example, neurons and heart muscle cells) rarely give rise to cancers. Figure 2.6.6.14.3.1 illustrates the most frequent occurrence of cancers in various body sites.

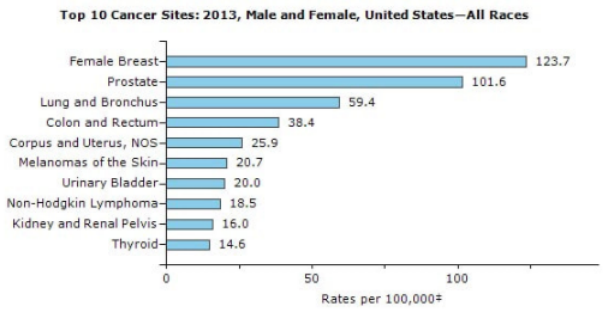


Figure 2.6.6.14.3.1. Top 10 Cancer Sites for Males and Females from All Races in the United States in 2013
(Image Source: CDC, **https://nccd.cdc.gov/uscs/toptencancers.aspx**)
‡ Rates are age-adjusted to the 2000 U.S. standard population (19 age groups – Census P25–1130)

While the prostate is the most common type of cancer that occurs in men, most survive with treatment. In contrast, other types of cancer are more often fatal. For example, the most common cancer, which causes death in men, even with treatment, is lung cancer. With women, a similar situation exists in that the breast is the most common site for cancer but more women die as a result of lung cancer.

What Do Cancers Look Like?

Cancer is a general term for more than 100 different cellular diseases, all with the same characteristic – the uncontrolled abnormal growth of cells in different parts of the body. Cancers appear in many forms. A few types are visible to the unaided eye but others grow inside the body and slowly destroy or replace internal tissues.

Skin Cancer

An example of a cancer that can be easily seen by the unaided eye is skin cancer. Skin cancers appear as raised, usually dark-colored, irregularly-shaped growths on the skin. As the cancer grows, it spreads to nearby areas of the skin. In advanced cases, the cancer metastasizes to lymph nodes and organs far away from the original site. The skin cancer illustrated in Figure 2.6.6.14.3.2 is known as a basal cell carcinoma. Melanomas and squamous cell carcinomas are other common skin cancers. Melanomas are usually the most malignant of the skin cancers.

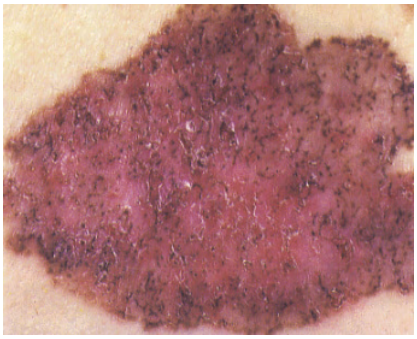


Figure 2.6.6.14.3.2. Photograph of basal cell carcinoma of the skin
(Image Source: NLM)

Other Cancers

Most cancers involve internal organs and require elaborate diagnostic tests to diagnose. Some large internal tumors can be felt or will push the skin outward and can be detected by noting abnormal bulges or an abnormal feel (for example, a hard area) to the body. Thyroid tumors, bone tumors, breast tumors, and testicular tumors are cancers that might be felt or observed by the patient. Other internal tumors may only be suspected based on diminished organ function (such as difficulty breathing with lung cancers), pain, bleeding (for example, blood in feces with colon cancer), weakness, or other unusual symptoms. To confirm the actual existence of a cancer may require diagnostic tests. This is especially the case where the cancer is not growing as a single large lump, but rather as a series of small tumors (metastatic foci) or when widely dispersed throughout the body (such as leukemia).
A few examples of internal cancers are presented in the following figures.

Liver Cancer

Numerous cancer nodules can be seen showing that much of the liver has been destroyed (Figure 2.6.6.14.3.3).

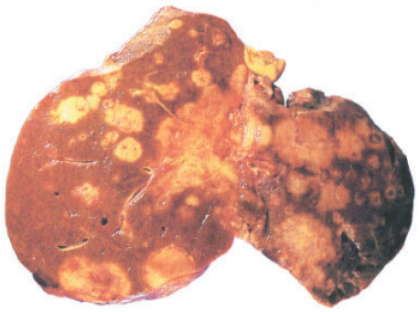


Figure 2.6.6.14.3.3. A liver with numerous cancer nodules
(Image Source: NLM)

Lung Cancer

An early developing squamous cell carcinoma can be seen growing in the middle of the lung (Figure 2.6.6.14.3.4). As the cancer develops, it will consume more of the lung and metastasize to other areas of the body.

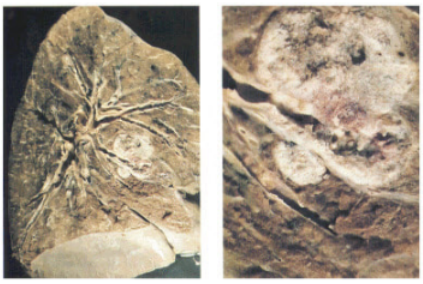


Figure 2.6.6.14.3.4. A cancerous lung
(Image Source: NLM)

Kidney Cancer

The photograph in Figure 2.6.6.14.3.5 shows the cancer has consumed much of the upper portion of the kidney.

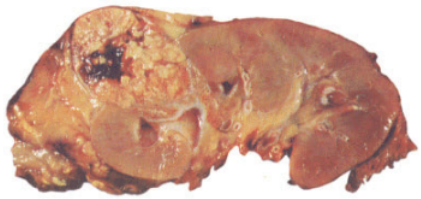


Figure 2.6.6.14.3.5. A kidney with cancer
(Image Source: NLM)

Historical Changes in Incidence of Cancer

Cancer has been recognized in humans for centuries. However, the incidence of various types of cancer has changed since the mid-1900s. This is especially true for lung and stomach cancer. Deaths from lung cancer hit a peak in the early 1990s and have been slowly declining since 2001. During that same period, deaths from stomach cancer decreased substantially. Breast cancer caused more deaths than any other type cancer in women for many decades. However, when women began smoking cigarettes, deaths from lung cancer outpaced deaths from breast cancer. These changes in types and incidences of cancer reflect the increased longevity of people as well as personal habits and environmental changes.

Latency Period for Cancer Development

Cancer is a chronic condition, which develops gradually over a period of time, and may become a clinical concern many years following the initial exposure to a carcinogen. This period of time is referred to as the **latency period**. The latency period varies with the type of cancers and may be as short as a few years to over 30 years. For example, the latency period for leukemia after benzene or radiation exposure may only be five years. In contrast, the latency period may be 20–30 years for skin cancer after arsenic exposure and mesothelioma (cancer of the pleura around the lungs) after asbestos exposure.

Survival Time

Success in treating cancer varies greatly with the type of cancer with some cancers responding to treatment whereas others do not. For example, medical treatment of cancers of the pancreas, liver, esophagus, and lung are largely unsuccessful. In contrast, cancers

of the thyroid, testes, and skin respond quite well to treatment. Table 2.6.6.14.3.3 shows the 5-year survival rate by cancer location.

Organ	Male 5-year Survival	Female 5-year Survival
Pancreas	7.6%	7.8%
Liver	17.1%	18.5%
Esophagus	18.3%	18.5%
Lung and Bronchus	14.9%	20.8%
Stomach	28.5%	33.2%
Brain and Nervous System	32.8%	35.1%
Leukemia	60.7%	58.3%
Ovary	----------	46.2%
Non-Hodgkin Lymphoma	69.2%	72.4%
Oral Cavity and Pharynx	65.0%	67.0%
Kidney	73.0%	75.0%
Colon/Rectum	64.9%	65.2%
Cervix	----------	67.5%
Hodgkin Lymphoma	85.4%	87.7%
Breast	83.6%	89.7%
Prostate	98.9%	----------
Urinary Bladder	78.9%	73.0%
Melanoma of the Skin	89.5%	94.0%
Testes	95.4%	----------
Thyroid	95.7%	98.7%
All Sites	67.0%	66.9%

Table 2.6.6.14.3.3. Five year survival rate by primary cancer site (2006-2012)
(Source: Table 1. Surveillance, Epidemiology, and End Results Program (SEER), National Cancer Institute, https://seer.cancer.gov/csr/1975_2013/results_merged/topic_survival.pdf)

What Causes Cancer?

A large number of industrial, pharmaceutical, and environmental chemicals have been identified as potential carcinogens by animal tests. Human epidemiology studies have confirmed that many are human carcinogens as well. However, while it is apparent that chemicals and radiation play a substantial role, it appears that lifestyle factors (such as diet, obesity, and smoking), and infections (such as hepatitis B, hepatitis C, and Human Papillomaviruses) are also major factors leading to the likelihood that a person will develop cancer. Additional factors that are involved in the development of cancer include aging and heredity.

Pathogenesis of Cancer

Carcinogenesis is a multi-step, multi-factorial genetic disease. All known tumors are composed of cells with genetic alterations that make them perform differently from their progenitor (parent) cells. The carcinogenesis process is very complex and unpredictable consisting of several phases and involving multiple genetic events (mutations) that take place over a very long period of time, at least 10 years for most types of cancer.

Cancer cells do not necessarily proliferate faster than their normal progenitors. In contrast to normal proliferating tissues where there is a strict and controlled balance between cell death and replacement, cancers grow and expand since more cancer cells are produced than die in a given time period. For a tumor to be detected it must attain a size of at least one cubic centimeter (about the size of a pea). This small tumor contains 100 million to a billion cells at that time. The development from a single cell to that size also means that the mass has doubled at least 30 times. During the long and active period of cell proliferation, the cancerous cells may have become aggressive in growth and have reverted to a less differentiated type cell that is not similar to the original cell

type.

While knowledge of carcinogenesis continues to evolve, it is clear that there are at least three main phases in cancer development:

1. Initiation
2. Promotion/Conversion
3. Progression

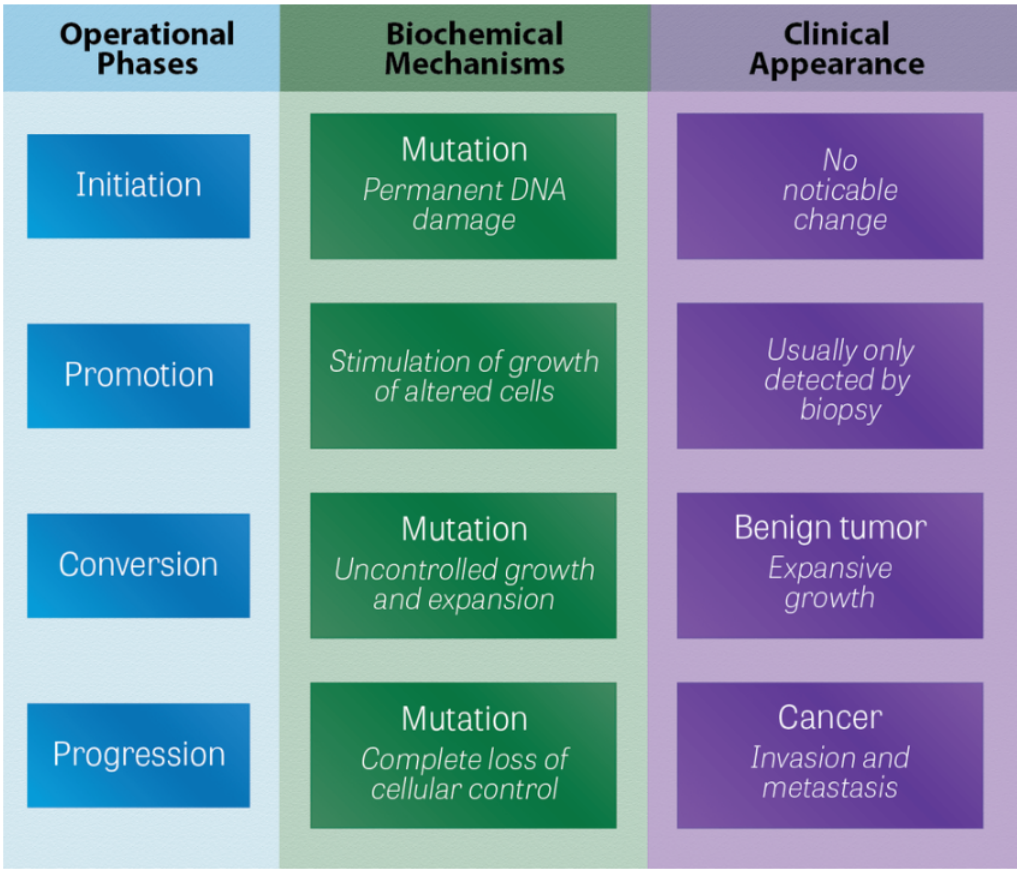


Figure 2.6.6.14.3.6. Phases of carcinogenesis
(Image Source: NLM)

1. Initiation

The **initiation** phase consists of the alteration of the DNA (mutation) of a normal cell, which is an irreversible change. The initiated cell has developed a capacity for individual growth. At this time, the initiated cell is indistinguishable from other similar cells in the tissue. The initiating event can consist of a single exposure to a carcinogenic agent or, in some cases, it may be an inherited genetic defect.

- An example is retinoblastoma in which some children who develop the disease may have inherited an altered copy of the gene involved and are at risk of passing the altered gene to successive generations.

The initiated cell, whether inherited or newly mutated, may remain dormant for months to years and, unless a promoting event occurs, may never develop into a clinical cancer case.

2. Promotion/Conversion

The **promotion/conversion** phase is the second major step in the carcinogenesis process in which specific agents (referred to as promoters) enhance the further development of the initiated cells. Promoters often, but not always, interact with the cell's DNA and

influence the further expression of the mutated DNA so that the initiated cell proliferates and progresses further through the carcinogenesis process. The clone of proliferating cells in this stage takes a form consistent with a benign tumor. The mass of cells remains as a cohesive group and physically keeps in contact with each other.

3. Progression

Progression is the third recognized step and is associated with the development of the initiated cell into a biologically malignant cell population. In this stage, a portion of the benign tumor cells may be converted into malignant forms so that a true cancer has evolved. Individual cells in this final stage can break away and start new clones of growth distant from the original site of development of the tumor. This is known as **metastasis**.

Genetic Activity

While the three-stage pathogenesis scheme describes the basic sequence of events in the carcinogenesis process, the actual events that take place in these various steps are due to activities of specific genes within the DNA of the cells. Cellular DNA contains two types of genes:

1. **Structural genes** direct the production of specific proteins within the cell.
2. **Regulatory genes** control the activity of the structural genes and direct the proliferation process of the cell.

The three classes of regulatory genes considered to have major roles in the carcinogenesis process are known as:

1. Proto-oncogenes
2. Oncogenes
3. Suppressor genes

Proto-oncogenes are normal cellular genes that encode and instruct the production of the regulatory proteins and growth factors within the cell or its membrane. The proteins encoded by proto-oncogenes are necessary for normal cellular cell growth and differentiation. Activation of a proto-oncogene can cause the alteration in the normal growth and differentiation of cells, which leads to neoplasia. Several agents can activate proto-oncogenes. This is the result of point mutations or by DNA rearrangements of the proto-oncogenes. The product of this proto-oncogene activation is an oncogene. Many proto-oncogenes have been identified and have usually been named after the source of their discovery, for example, the KRAS proto-oncogene was named for the discovery using the Kirsten rat sarcoma virus. HRAS, MYC, MYB, and SRC are other examples of proto-oncogenes. The proto-oncogenes are not specific for the original species but have been found in many other species, including humans. These proto-oncogenes are present in many cells but remain dormant until activated. Either a point mutation or chromosomal damage of various types can induce activation. Once activated they become an oncogene.

Oncogenes are altered or misdirected proto-oncogenes which now have the ability to direct the production of proteins within the cell that can change or transform the normal cell into a neoplastic cell. Most oncogenes differ from their proto-oncogenes by a single point mutation located at a specific codon (a group of three DNA bases that encodes for a specific amino acid) of a chromosome. The altered DNA in the oncogene results in the production of an abnormal protein that can alter cell growth and differentiation. It appears that a single activated oncogene is not sufficient for the growth and progression of a cell and its offspring to form a cancerous growth. However, it is a major step in the carcinogenesis process.

Tumor suppressor genes, sometimes referred to as anti-oncogenes, are present in normal cells and serve to counteract and change the proto-oncogenes and altered proteins that they are responsible for. The tumor suppressor genes serve to prevent a cell with damaged DNA from proliferating and evolving into an uncontrolled growth. They actively function to effectively oppose the action of an oncogene. If a tumor suppressor gene is inactivated (usually by a point mutation), its control over the oncogene and transformed cell may be lost. Thus the tumor-potential cell can now grow without restraint and is free of the normal cellular regulatory control. The suppressor gene most frequently altered in human tumors is the p53 gene. Damaged p53 genes have been identified in over 50% of human cancers.

The **p53 gene** normally halts cell division and stimulates repair enzymes to rebuild and restore the damaged regions of the DNA. If

the damage is too extensive, the p53 commands the cell to self-destruct. An altered p53 is incapable of these defensive actions and cannot prevent the cell with damaged DNA from dividing and proliferating in an erratic and uncontrolled manner. This is the essence of cancer.

This section represents only a brief overview of an enormously complex process for which knowledge is continuously evolving with the tools of molecular biology. New factors are continuously being identified; however, many pieces of this cancer puzzle remain elusive at this time.

Knowledge Check

1) A body growth with the ability to metastasize or invade into surrounding tissues is known as a:

a) Benign tumor
b) Malignant tumor
c) Hyperplasia

Answer

Malignant tumor - **This is the correct answer.**
A malignant tumor that has the ability to metastasize or invade into surrounding tissues. It is the same as cancer.

2) Most cancers are thought to be due to the following:

a) Infections
b) Food additives
c) Lifestyle factors
d) Pollution

Answer

Lifestyle factors - **This is the correct answer.**
Lifestyle (including diet, tobacco use, reproductive and sexual behavior, and alcohol consumption) is considered to cause about 75% of all cancers.

3) The initial stage in carcinogenesis in which there is an alteration of the DNA (mutation) is referred to as the:

a) Progression stage
b) Promotion stage
c) Initiation stage

Answer

Initiation stage - **This is the correct answer.**
The initiation phase consists of the alteration of the DNA (mutation) of a normal cell, which is irreversible change.

4) The cellular gene which is present in most normal cells and serves as a balance to the genes for tumor expression is known as a:

a) Tumor suppressor gene
b) Oncogene
c) Proto-oncogene

Answer

Tumor suppressor gene - **This is the correct answer.**
Tumor suppressor genes, sometimes referred to as anti-oncogenes, are present in normal cells and serve as a balance to the genes for expression or proto-oncogenes.

This page titled 2.6.6.14.3: Cancer is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **14.3: Cancer** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

Neurotoxicity

The nervous system is very complex and toxins can act at many different points in this complex system. The focus of this section is to provide a basic overview of how the nervous system works and how neurotoxins affect it. Due to the complexity of these topics, this section does not include extensive details related to the anatomy and physiology of the nervous system or the many neurotoxins in our environment and the subtle ways they can damage the nervous system or interfere with its functions.

Since the nervous system innervates all areas of the body, some toxic effects may be quite specific and others generalized depending upon where in the nervous system the toxin exerts its effect. Before discussing how neurotoxins cause damage, we will look at the basic anatomy and physiology of the nervous system.

Anatomy and Physiology of the Nervous System

The nervous system has three basic functions:

1. Specialized cells detect sensory information from the environment and relay that information to other parts of the nervous system.
2. It directs motor functions of the body usually in response to sensory input.
3. It integrates the thought processes, learning, and memory.

All of these functions are potentially vulnerable to the actions of toxicants.
The nervous system consists of two fundamental anatomical divisions:

1. Central nervous system (CNS)
2. Peripheral nervous system (PNS)

Central Nervous System

The CNS includes the brain and spinal cord. The CNS serves as the control center and processes and analyzes information received from sensory receptors and in response issues motor commands to control body functions. The brain, which is the most complex organ of the body, structurally consists of six primary areas (Figure 2.6.6.14.4.1):

1. **Cerebrum** — controls thought processes, intelligence, memory, sensations, and complex motor functions.
2. **Diencephalon *(thalamus, hypothalamus, pituitary gland)*** — relays and processes sensory information; controlls emotions, autonomic functions, and hormone production.
3. **Midbrain** — processes auditory and visual data; generates involuntary motor responses.
4. **Pons** — a tract and relay center which also assists in somatic and visceral motor control.
5. **Cerebellum** — voluntary and involuntary motor activities based on memory and sensory input.
6. **Medulla oblongata** — relays sensory information to the rest of the brain; regulates autonomic function, including heart rate and respiration.

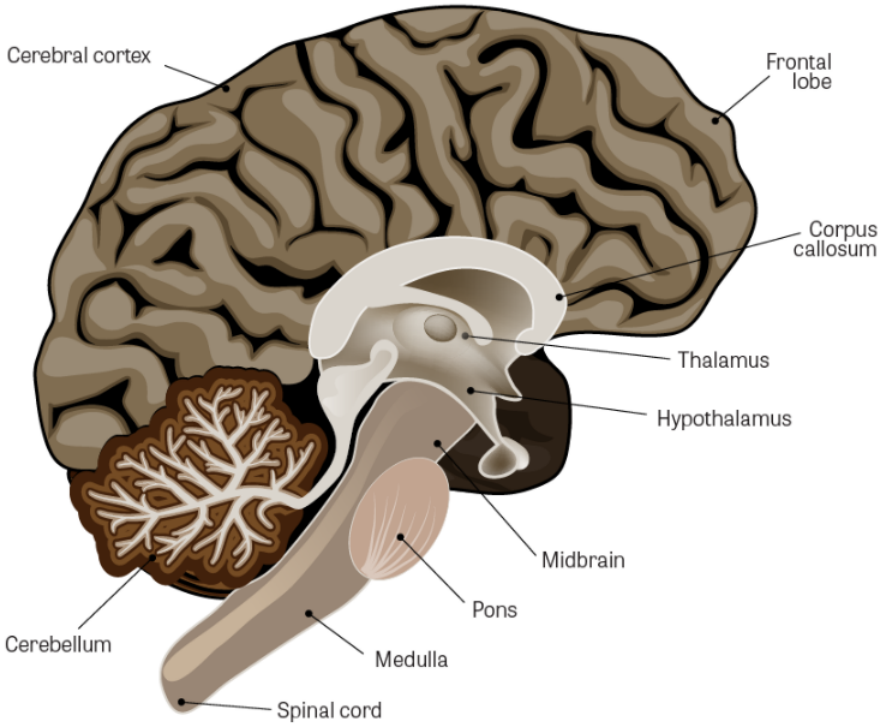


Figure 2.6.6.14.4.1. Internal anatomy of the brain
(Image Source: Adapted from iStock Photos, ©)

Peripheral Nervous System

The PNS consists of all nervous tissue outside the CNS (Figure 2.6.6.14.4.2). The PNS contains two forms of nerves:

1. Afferent nerves, which relay sensory information to the CNS.
2. Efferent nerves, which relay motor commands from the CNS to various muscles and glands.

Efferent nerves are organized into two systems. One is the **somatic nervous system** that is also known as the voluntary system and which carries motor information to skeletal muscles. The second efferent system is the **autonomic nervous system**, which carries motor information to smooth muscles, cardiac muscle, and various glands. The major difference between these two systems pertains to conscious control.

- The somatic system is under our voluntary control such as moving our arms by consciously telling our muscles to contract.
- In contrast, we cannot consciously control the smooth muscles of the intestine, heart muscle, or secretion of hormones. Those functions are automatic and involuntary as controlled by the autonomic nervous system.

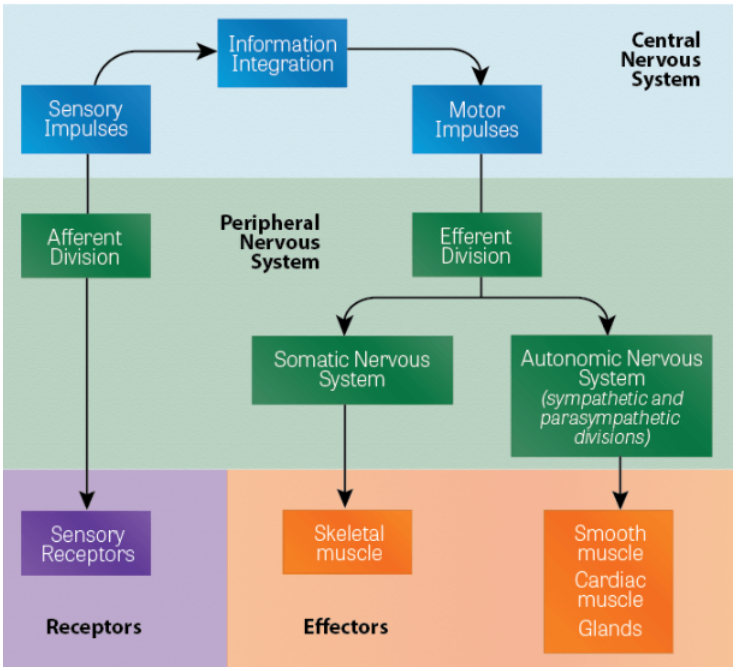


Figure 2.6.6.14.4.2. Structures of the central nervous system and peripheral nervous system (Image Source: NLM)

Cells of the Nervous System

There are two categories of cells found in the nervous system: neurons and glial cells. **Neurons** are the functional nerve cells directly responsible for transmission of information to and from the CNS to other areas of the body. **Glial cells** (also known as neuroglia) provide support to the neural tissue, regulate the environment around the neurons, and protect against foreign invaders.

Neurons communicate with all areas of the body and are present within both the CNS and PNS. They serve to transmit rapid impulses to and from the brain and spinal cord to virtually all tissues and organs of the body. As such, they are an essential cell and their damage or death can have critical effects on body function and survival. When neurons die, they are not replaced. As neurons are lost, so are certain neural functions such as memory, ability to think, quick reactions, coordination, muscular strength, and our various senses such as sight, hearing, and taste. If the neuron loss or impairment is substantial, severe and permanent disorders can occur, such as blindness, paralysis, and death.

A neuron consists of a cell body and two types of extensions, numerous dendrites, and a single axon (Figure 2.6.6.14.4.3). **Dendrites** are specialized in receiving incoming information and sending it to the neuron cell body with transmission (electrical charge) on down the axon to one or more junctions with other neurons or muscle cells (known as synapses). The **axon** may extend long distances, over a meter in some cases, to transmit information from one part of the body to another. The **myelin sheath** is a multi-layer coating that wraps some axons and helps insulate the axon from surrounding tissues and fluids, and prevents the electrical charge from escaping from the axon.

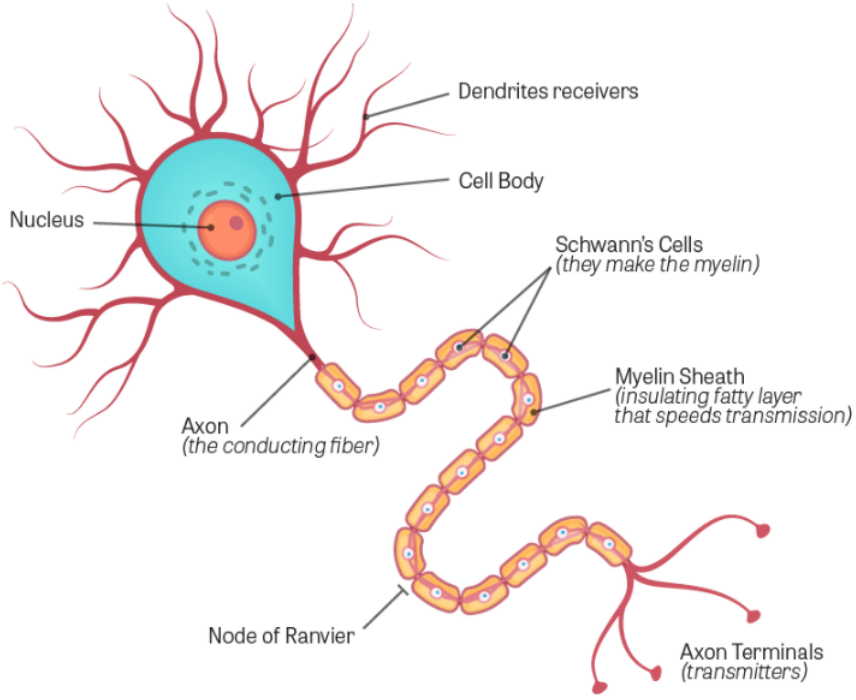


Figure 2.6.6.14.4.3. Neuron structure
(Image Source: Adapted from iStock Photos, ©)

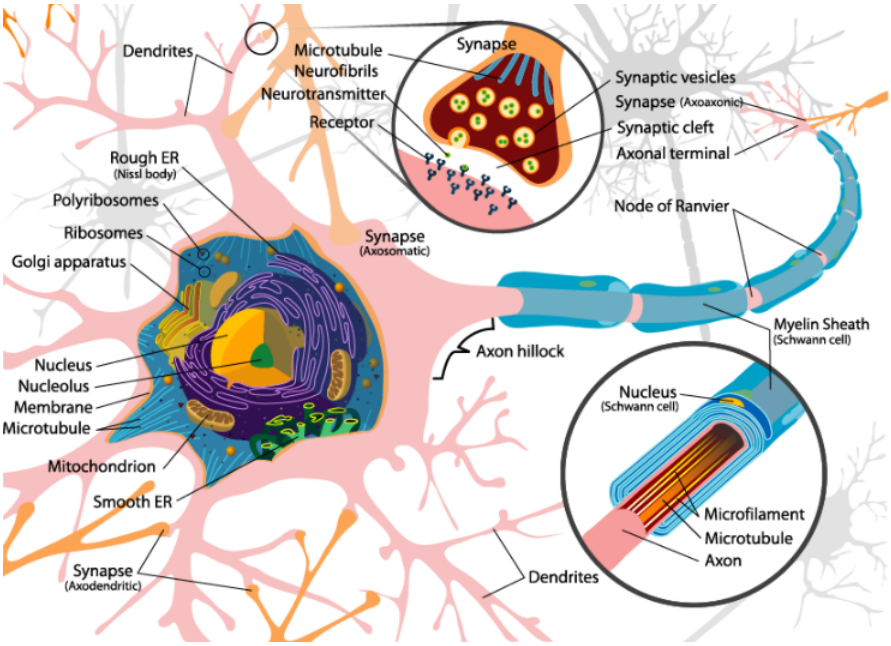


Figure 2.6.6.14.4.4. Complete neuron cell diagram
(Image Source: Adapted from Wikimedia Commons, obtained under Public Domain. Author: LadyofHats.)

Information passes along the network of neurons between the CNS and the sensory receptors and the effectors by a combination of electrical pulses and chemical neurotransmitters. The information (electrical charge) moves from the dendrites through the cell body and down the axon. The mechanism by which an electrical impulse moves down the neuron is quite complex. When the neuron is at rest, it has a negative internal electrical potential. This changes when a neurotransmitter binds to a dendrite receptor. Protein channels of the dendrite membrane open allowing the movement of charged chemicals across the membrane, which creates

an electrical charge. The propagation of an electrical impulse (known as action potential) proceeds down the axon by a continuous series of openings and closings of sodium-potassium channels and pumps. The action potential moves like a wave from one end (dendritic end) to the terminal end of the axon.

However, the electrical charge cannot cross the gap (synapse) between the axon of one neuron and the dendrite of another neuron or an axon and a connection with a muscle cell (neuromuscular junction). Chemicals called neurotransmitters move the information across the synapse.

Neurons do not make actual contact with one another but have a gap, known as a **synapse**. As the electrical pulse proceeds up or down an axon, it encounters at least one junction or synapse. An electrical pulse cannot pass across the synapse. At the terminal end of an axon is a synaptic knob, which contains the neurotransmitters.

Neurotransmitters

Vesicles release **neurotransmitters** upon stimulus by an impulse moving down the presynaptic neuron. The neurotransmitters diffuse across the synaptic junction and bind to receptors on the postsynaptic membrane. The neurotransmitter-receptor complex then initiates the generation of an impulse on the next neuron or the effector cell, for example, a muscle cell or secretory cell.

After the impulse has again been initiated, the neurotransmitter-complex must be inactivated or continuous impulses (beyond the original impulse) will be generated. Enzymes perform this inactivation, which serves to break down the complex at precisely the right time and after the exact impulse has been generated. There are several types of neurotransmitters and corresponding inactivating enzymes. One of the major neurotransmitters is acetylcholine with acetylcholinesterase as the specific inactivator.

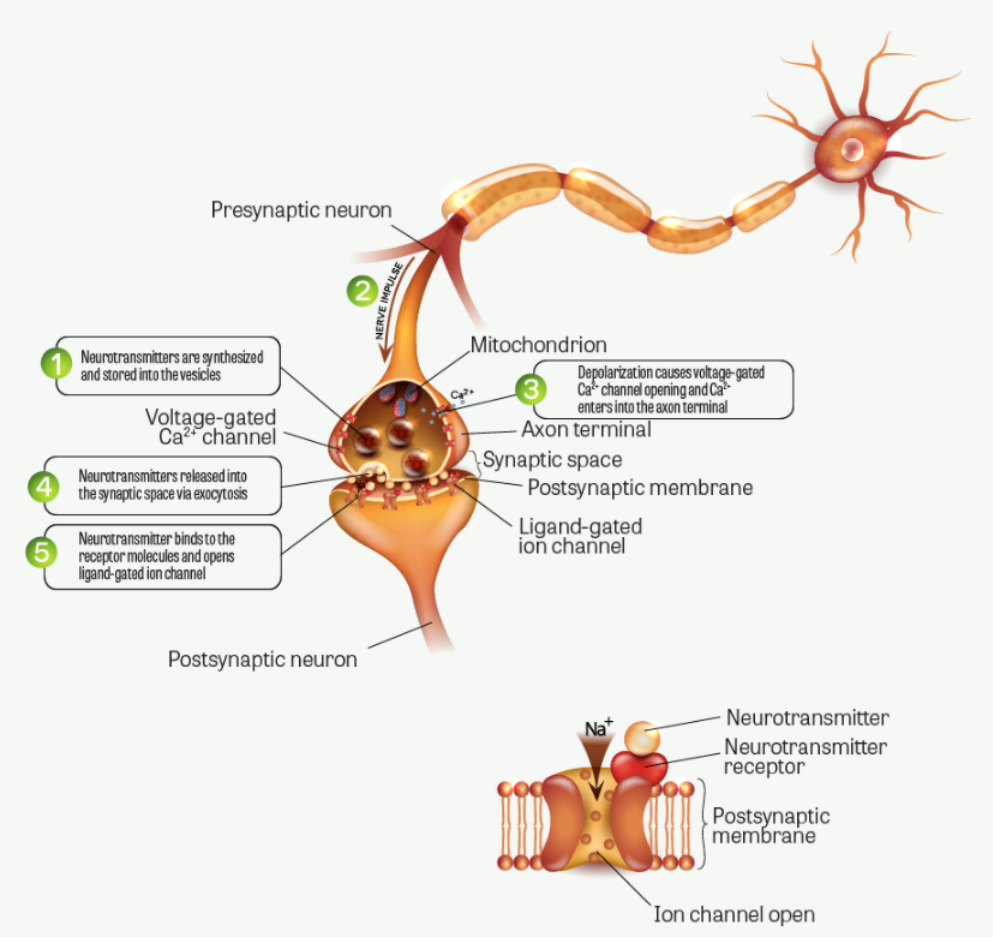


Figure 2.6.6.14.4.5. Impulse transmission across synapse
(Image Source: Adapted from iStock Photos, ©)

There are over 100 known neurotransmitters. Among the most well-known are:

- Acetylcholine
- Dopamine
- Serotonin
- Norepinephrine
- GABA (gamma-aminobutyric acid)

Types of Neurons

Neurons are categorized by their function and consist of three types:

1. **Sensory neurons (afferent neurons)** carry information from sensory receptors (usually processes of the neuron) to the CNS. Some sensory receptors detect external changes such as temperature, pressure, and the senses of touching and vision. Others monitor internal changes such as balance, muscle position, taste, deep pressure, and pain.
2. **Motor neurons (effector neurons)** relay information from the CNS to other organs terminating at the effectors. Motor neurons make up the efferent neurons of both the somatic and autonomic nervous systems.
3. **Interneurons (association neurons)** are located only in the CNS and provide connections between sensory and motor neurons. They can carry either sensory or motor impulses. They are involved in spinal reflexes, analysis of sensory input, and coordination of motor impulses. They also play a major role in memory and the ability to think and learn.

Glial Cells

Glial cells are important as they provide a structure for the neurons by protecting them from outside invading organisms, and maintaining a favorable environment (nutrients, oxygen supply, etc.). The neurons are highly specialized and do not have all the usual cellular organelles to provide them with the same life-support capability. They are highly dependent on the glial cells for their survival and function. For example, neurons have such a limited storage capacity for oxygen that they are extremely sensitive to decreases in oxygen (anoxia) and will die within a few minutes. The list below describes the types of glial cells:

- **Astrocytes** are big cells, only in the CNS, and maintain the blood-brain barrier that controls the entry of fluid and substances from the circulatory system into the CNS. They also provide rigidity to the brain structure.
- **Schwann cells and oligodendrocytes** wrap themselves around some axons to form **myelin**, which serves like insulation. Myelinated neurons usually transmit impulses at high speed, such as needed in motor neurons. Loss of myelination causes a dysfunction of these cells.
- **Microglia** are small, mobile, phagocytic cells.
- **Ependymal cells** produce the cerebrospinal fluid (CSF) which surrounds and cushions the central nervous system.

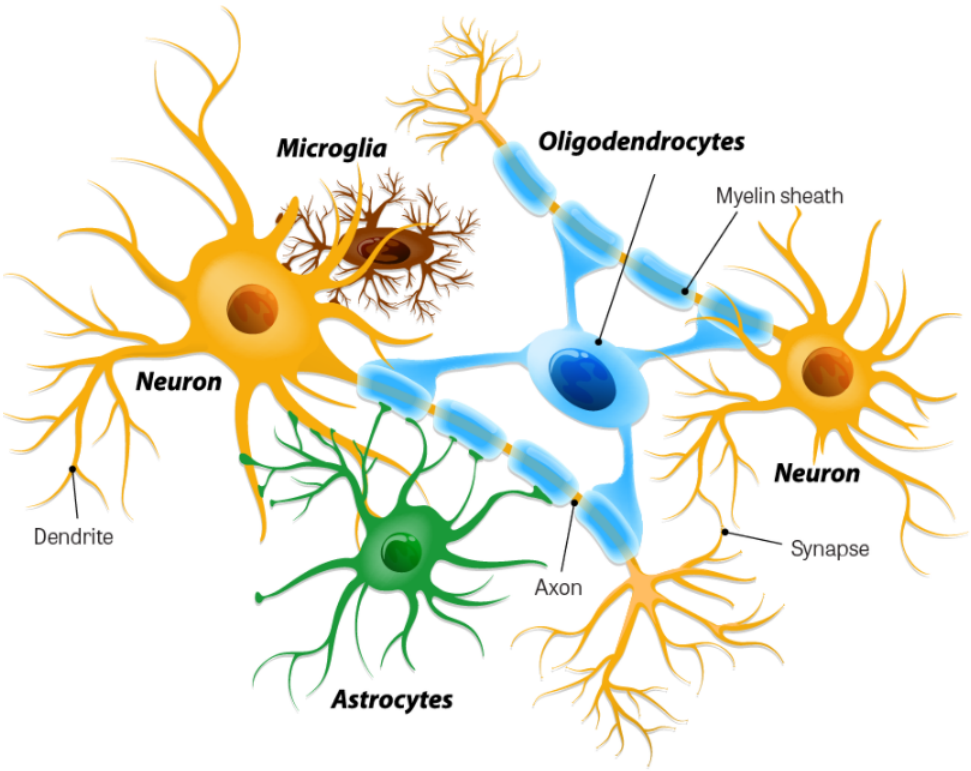


Figure 2.6.6.14.4.6. Neurons and neuroglial cells
(Image Source: Adapted from iStock Photos, ©)

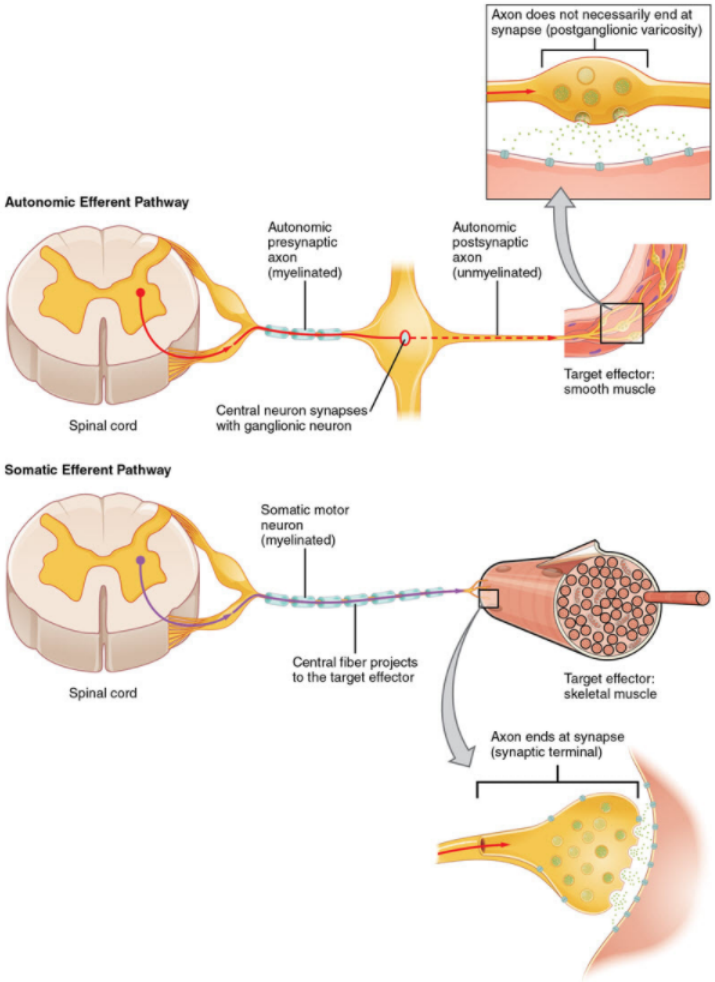


Figure 2.6.6.14.4.7. Comparison of somatic and visceral reflects
(Image Source: Wikimedia Commons, obtained under Creative Commons Attribution 3.0 Unported License. Author: OpenStax College. View original image. Source: Anatomy & Physiology, Connexions Web site. http://cnx.org/content/col11496/1.6/, Jun 19, 2013.)

Toxic Damage to Nervous System

The nervous system is quite vulnerable to toxins since chemicals interacting with neurons can change the critical voltages, which must be carefully maintained. However, the nervous system has defense mechanisms that can protect it from toxins.

Most of the CNS is protected by an anatomical barrier between the neurons and blood vessels, known as the **blood-brain barrier**. It is protected from some toxin exposures by tightening junctions between endothelial cells of the blood vessels in the CNS and having astrocytes surround the blood vessels. This prevents the diffusion of chemicals out of the blood vessels and into the intracellular fluid except for small, lipid-soluble, non-polar molecules. Specific transport mechanisms exist to transport essential nutrients (such as glucose and amino acids and ions) into the brain. Another defense mechanism within the brain to counter chemicals that pass through the vascular barrier is the presence of metabolizing enzymes. Certain detoxifying enzymes, such as monoamine oxidase, can biotransform many chemicals to less toxic forms as soon as they enter the intercellular fluid.

The basic types of changes due to toxins can be divided into three categories – 1) sensory; 2) motor; and 3) interneuronal – depending on the type of damage sustained.

1. Damage can occur to sensory receptors and **sensory neurons**, which can affect the basic senses of pressure, temperature, vision, hearing, taste, smell, touch, and pain.
    - For example, heavy metal poisoning (especially lead and mercury) can cause deafness and loss of vision.
    - Several chemicals including inorganic salts and organophosphorus compounds can cause a loss of sensory functions.
2. Damage to **motor neurons** can cause muscular weakness and paralysis.
    - Isonicotinic hydrazide (used to treat tuberculosis) can cause such damage.
3. **Interneuronal** damage can cause learning deficiencies, loss of memory, incoordination, and emotional conditions.
    - Low levels of inorganic mercury and carbon monoxide can cause depression and loss of memory.

Mechanisms for Toxic Damage to the Nervous System

Toxic damage to the nervous system occurs by the following basic mechanisms:

1. Direct damage and death of neurons and glial cells.
2. Interference with electrical transmission.
3. Interference with chemical neurotransmission.

A. Death of Neurons and Glial Cells

The most common cause of death of neurons and glial cells is **anoxia**, an inadequate oxygen supply to the cells or their inability to utilize oxygen. Anoxia may result from the blood's decreased ability to provide oxygen to the tissues (impaired hemoglobin or decreased circulation) or from the cells unable to utilize oxygen.

- For example, carbon monoxide and sodium nitrite can bind to hemoglobin preventing the blood from being able to transport oxygen to the tissues.
- Hydrogen cyanide and hydrogen sulfide can penetrate the blood-brain barrier and is rapidly taken up by neurons and glial cells.
- Another example is sodium fluoroacetate (commonly known as Compound 1080, a rodent pesticide) which inhibits a cellular enzyme.

Those chemicals interfere with cellular metabolism and prevent nerve cells from being able to utilize oxygen. This is called **histoxic anoxia**.
Neurons are among the most sensitive cells in the body to inadequate oxygenation. Lowered oxygen for only a few minutes is sufficient to cause irreparable changes leading to the death of neurons.
Several other neurotoxins directly damage or kill neurons, including:

- Lead
- Mercury
- Some halogenated industrial solvents including methanol (wood alcohol)
- Toluene
- Trimethyltin polybrominated diphenyl ethers (PBDEs)

While some neurotoxic agents affect neurons throughout the body, others are quite selective.

- For example, methanol specifically affects the optic nerve, retina, and related ganglion cells while trimethyltin kills neurons in the hippocampus, a region of the cerebrum.

Other agents can degrade neuronal cell function by diminishing its ability to synthesize protein, which is required for the normal function of the neuron.

- Organomercury compounds exert their toxic effect in this manner.

With some toxins, only a portion of the neuron is affected. If the cell body is killed, the entire neuron will die. Some toxins can cause death or loss of only a portion of the dendrites or axon while the cell itself survives but with diminished or total loss of function. Commonly axons begin to die at the very distal end of the axon with necrosis slowly progressing toward the cell body. This is referred to as "dying-back neuropathy."

- Some organophosphate chemicals (including some pesticides) cause this distal axonopathy. The mechanism for the dying back is not clear but may be related to the inhibition of an enzyme (neurotoxic esterase) within the axon.
- Other well-known chemicals can cause distal axonopathy include ethanol, carbon disulfide, arsenic, ethylene glycol (in antifreeze), and acrylamide.

B. Interference with Electrical Transmission

There are two basic ways that a foreign chemical can interrupt or interfere with the propagation of the electrical potential (impulse) down the axon to the synaptic junction:

1. To interfere with the movement of the action potential down the intact axon.
2. To cause structural damage to the axon or its myelin coating. Without an intact axon, transmission of the electrical potential is not possible.

Agents that can block or interfere with the sodium and potassium channels and sodium-potassium pump cause interruption of the propagation of the electrical potential. This will weaken, slow, or completely interrupt the movement of the electrical potential. Many potent neurotoxins exert their toxicity by this mechanism.

- Tetrodotoxin (a toxin in frogs, pufferfish, and other invertebrates) and saxitoxin (a cause of shellfish poisoning) blocks sodium channels. Batrachotoxin (a toxin in South American frogs used as arrow poison) and some pesticides (DDT and pyrethroids) increases the permeability of the neuron membrane preventing closure of sodium channels which leads to repetitive firing of the electrical charge and an exaggerated impulse.

A number of chemicals can cause **demyelination**. Many axons (especially in the PNS) are wrapped with a protective myelin sheath that acts as insulation and restricts the electrical impulse within the axon. Agents that selectively damage these coverings disrupt or interrupt the conduction of high-speed neuronal impulses. Loss of a portion of the myelin can allow the electrical impulse to leak out into the tissue surrounding the neuron so that the pulse does not reach the synapse with the intended intensity.

- In some diseases, such as Multiple Sclerosis (MS) and Amyotrophic Lateral Sclerosis (ALS), the myelin is lost, causing paralysis and loss of sensory and motor function.

A number of chemicals can cause demyelination:

- Diphtheria toxin causes loss of myelin by interfering with the production of protein by the Schwann cells that produce and maintain myelin in the PNS.
- Triethyltin (used as a biocide, preservative, and polymer stabilizer) interrupts the myelin sheath around peripheral nerves.
- Lead causes loss of myelin primarily around peripheral motor axons.

C. Interference with Chemical Neurotransmission

Synaptic dysfunction is a common mechanism for the toxicity of a wide variety of chemicals. There are two types of synapses: *those between two neurons* (axon of one neuron and dendrites of another) and those *between a neuron and a muscle cell or gland*. The basic mechanism for the chemical transmission is the same. The major difference is that the neurotransmitting chemical between a neuron and muscle cell is acetylcholine whereas there are several other types of neurotransmitting chemicals involved between neurons, depending on where in the nervous system the synapse is located.
There are four basic steps involved in neurotransmission at the synapse:

1. Synthesis and storage of neurotransmitter (synaptic knob of axon).
2. Release of the neurotransmitter (synaptic knob with movement across synaptic cleft).
3. Receptor activation (effector membrane).
4. Inactivation of the transmitter (enzyme breaks down neurotransmitter stopping induction of action potential).

The arrival of the action potential at the synaptic knob initiates a series of events culminating in the release of the chemical neurotransmitter from its storage depots in vesicles. After the neurotransmitter diffuses across the synaptic cleft, it complexes with a receptor (membrane-bound macromolecule) on the post-synaptic side. This binding causes an ion channel to open, changing the membrane potential of the post-synaptic neuron or muscle or gland. This starts the process of impulse formation or action potential

in the next neuron or receptor cell. However, unless this receptor-transmitter complex is inactivated, the channel remains open with continued pulsing. Thus, the transmitter action must be terminated. Specific enzymes that can break the bond and return the receptor-membrane to its resting state do this.

Drugs and environmental chemicals can interact at specific points in this process to change the neurotransmission. Depending on where and how the xenobiotics act, the result may be either an increase or a decrease in neurotransmission. Many drugs (such as tranquilizers, sedatives, stimulants, beta-blockers) are used to correct imbalances to neurotransmissions (such as occurs in depression, anxiety, and cardiac muscular weakness). The mode of action of some analgesics is to block receptors, which prevent transmission of pain sensations to the brain.

Exposure to environmental chemicals that can perturb neurotransmission is a very important area of toxicology. Generally, neurotoxins affecting neurotransmission act to:

1. Increase or decrease the release of a neurotransmitter at the presynaptic membrane.
2. Block receptors at the postsynaptic membrane.
3. Modify the inactivation of the neurotransmitter.

This is a list of only a few examples of neurotoxins to show the range of mechanisms:

- α-Bungarotoxin (a potent venom of elapid snakes) prevents the release of neurotransmitters.
- Scorpion venom potentiates the release of a neurotransmitter (acetylcholine).
- Black widow spider venom causes an explosive release of neurotransmitters.
- Botulinum toxin blocks the release of acetylcholine at neuromuscular junctions.
- Atropine blocks acetylcholine receptors.
- Strychnine inhibits the neurotransmitter glycine at postsynaptic sites resulting in an increased level of neuronal excitability in the CNS.
- Nicotine binds to certain cholinergic receptors.

A particularly important type of neurotoxicity is the inhibition of acetylcholinesterase. The specific function of acetylcholinesterase is to stop the action of acetylcholine once it has bound to a receptor and initiated the action potential in the second nerve or at the neuro-muscular or glandular junction. If the acetylcholine-receptor complex is not inactivated, continual stimulation will result leading to paralysis and death.

- Many commonly used chemicals, especially organophosphate and carbamate pesticides, poison mammals by this mechanism.
- The major military nerve gases are also cholinesterase inhibitors.

Acetylcholine is a common neurotransmitter. It is responsible for transmission at all neuromuscular and glandular junctions as well as many synapses within the CNS.

Events Involved in a Typical Cholinergic SynapseThe complexity of the sequence of events that takes place at a typical cholinergic synapse is indicated below:

Step	Events
1	• Electrical impulse arrives at synaptic knob and depolarizes the bulb and presynaptic membrane. • Synaptic vesicles release acetylcholine *(ACh)*.
2	• Calcium ions enter the synaptic knob's cytoplasm. • Synaptic vesicles release ACh.
3	• ACh release stops as calcium ions are removed from the synaptic knob cytoplasm. • The released ACh diffuses across synaptic cleft and binds to receptors on the post-synaptic membrane. • The chemically-regulated receptors cause a graded depolarization on the postsynaptic surface, which then is transmitted on down the axon or into the effector cell.
4	• ACh is broken down by acetylcholinesterase into choline and acetate at the receptor site on the postsynaptic membrane. • Choline is then reabsorbed from the synaptic cleft and is available for resynthesis into more ACh and stored by the synaptic vesicles for future use.

Table 1. Events that take place at a typical cholinergic synapse

The nervous system is the most complex system of the body. There are still many gaps in understanding how many neurotoxins act, yet research is discovering their possible effects on the body's structures and functions. It is important to understand that the most potent toxins (on a weight basis) are neurotoxins with extremely minute amounts sufficient to cause death.

Knowledge Check

1) The two fundamental anatomical divisions of the nervous system are the:

a) Cerebrum and cerebellum

b) Central nervous system and peripheral nervous system

c) Brain and spinal cord

Answer

Central nervous system and peripheral nervous system - **This is the correct answer.**

The two fundamental anatomical divisions of the nervous system are the Central Nervous System (brain and spinal cord) and the Peripheral Nervous System, which consists of all nerves outside the brain and spinal cord.

2) The two major categories of cells found in the nervous system are:

a) Neurons and glial cells

b) Astrocytes and microglia

c) Schwann cells and oligodendrocytes

Answer

Neurons and glial cells - **This is the correct answer.**

The two major categories of cells found in the nervous system are neurons and glial cells. Neurons are the functional nerve cells directly responsible for transmission of information to and from the CNS to other areas of the body. Glial cells (also known as neuroglia) provide support to the neural tissue, regulate the environment around the neurons, and protect against foreign invaders.

3) The propagation of an electrical impulse (action potential) down an axon consists of:

a) The transmission of the action potential by chemical neurotransmitters
b) The movement of sodium ions from the dendrite to the axon
c) A continuous series of opening and closing of sodium-potassium channels and pumps

Answer

A continuous series of opening and closing of sodium-potassium channels and pumps - **This is the correct answer.**
The propagation of an electrical impulse (action potential) down an axon consists of a continuous series of opening and closing of sodium-potassium channels and pumps. The action potential moves like a wave from one end (dendritic end) to the terminal end of the axon.

4) The type of neuron that relays information from the CNS to other organs is a:

a) Motor neuron
b) Sensory neuron
c) Interneuron

Answer

Motor neuron - **This is the correct answer.**
Motor Neurons (effector neurons) relay information from the CNS to other organs terminating at the effectors.

5) The primary cause of death to neurons and glial cells is:

a) Interference with chemical transmission
b) Interference with electrical transmission
c) Anoxia

Answer

Anoxia - **This is the correct answer.**
The most common cause of death of neurons and glial cells is anoxia, an inadequate oxygen supply to the cells or their inability to utilize oxygen.

6) A major mechanism that prevents the action potential (impulse) from moving down an axon is:

a) Blockage or interference with movement of sodium and potassium ions in and out of neuron membrane, changing the action potential
b) Excessive release of chemical neurotransmitters
c) Blocking receptors at the post-synaptic membrane

Answer

Blockage or interference with movement of sodium and potassium ions in and out of neuron membrane, changing the action potential - **This is the correct answer.**
Interruption of the propagation of the electrical potential is caused by agents that can block or interfere with the sodium and potassium channels and sodium-potassium pump. This will weaken, slow, or completely interrupt the movement of the electrical potential.

7) What are the two basic types of synapses?

a) Neuro-muscular and neuro-glandular

b) CNS and PNS synapses
c) Those between two neurons and a neuron and effector

Answer

Those between two neurons and a neuron and effector - **This is the correct answer.**
The two basic types of synapses are those between two neurons and those between a neuron and effectors, such as muscle cell or gland. The major difference in the two basic types is that the neurotransmitting chemical between a neuron and muscle cell is acetylcholine whereas there are several other types of neurotransmitting chemicals involved between neurons, depending on where in the nervous system the synapse is located.

This page titled 2.6.6.14.4: Neurotoxicity is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **14.4: Neurotoxicity** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.15: Intuitive Toxicology and Risk Communication

Learning Objectives

After completing this lesson, you will be able to:

- Describe aspects of Intuitive Toxicology.
- Describe aspects of Risk Communication.

In this section...

Topics include:

2.6.6.15.1: Intuitive Toxicology

2.6.6.15.2: Risk Communication

Section 15: Key Points

What We've Covered

This section made the following main points:

- Intuitive toxicology:
  - Combines the intuitive elements of expert and public risk judgments involved with exposure assessment, toxicology, and risk assessment.
  - Asks if toxicologists agree or disagree that an animal's reactions to a chemical reliably predict human reactions.
  - Asks if a chemical found in a scientific study to cause cancer in animals can be reasonably assumed to cause cancer in humans.
- When communicating risk, one needs to:
  - Accept and involve the public as a legitimate partner.
  - Listen to the public’s specific concerns.
  - Be honest, frank, and open.
  - Coordinate and collaborate with other credible sources.
  - Meet the needs of the media.
  - Speak clearly and with compassion.
  - Plan carefully and evaluate your efforts.
  - Acknowledge and describe uncertainty, such as any data gaps or issues relating to methodology.

This page titled 2.6.6.15: Intuitive Toxicology and Risk Communication is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.6.15.1: Intuitive Toxicology

Intuitive Toxicology

Intuitive can be defined as "using or based on what one feels to be true even without conscious reasoning." Humans have always been intuitive toxicologists via the use of the senses of sight, taste, and smell to try to detect harmful or unsafe food, water, and air.

Intuition and Professional Judgment in Toxicology

Even well-established scientific approaches used in human risk assessment depend on extrapolations and judgments when assessing human, animal and other toxicology data. This led to the study of **intuitive toxicology**—the intuitive elements of expert and public risk judgments involved with exposure assessment, toxicology, and risk assessment.

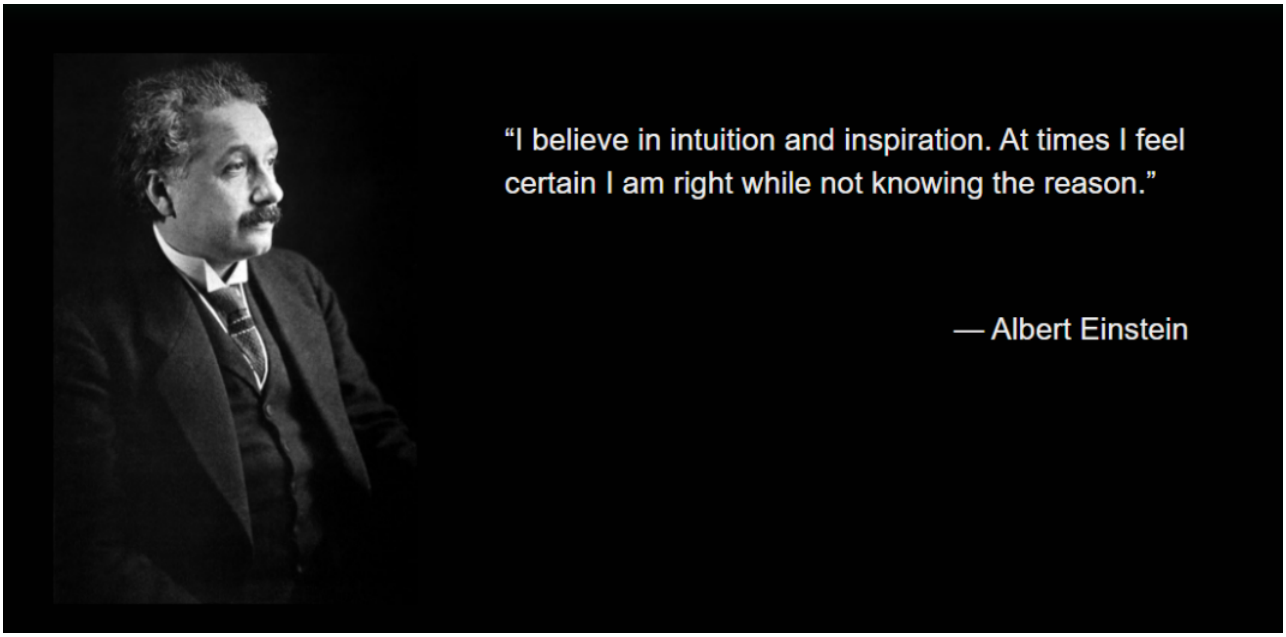


Figure 2.6.6.15.1.1. Albert Einstein understood the importance of intuition along with knowledge and experience (Image Source: Wikimedia Commons, Public Domain - original image)

Studies of Intuitive Toxicology

The studies of intuitive toxicology have surveyed toxicologists (for example, members of the Society of Toxicology) and others about a wide range of attitudes, beliefs, and perceptions regarding risks from chemicals. These have included basic concepts, assumptions, and interpretations related to the effects of chemical concentration, dose, and exposure on risk, and the value of animal studies for predicting the effects of chemicals on humans.

Two questions that have been studied repeatedly in intuitive toxicology are:

1. Would you agree or disagree that the way an animal reacts to a chemical is a reliable predictor of how a human would react to it?
2. If a scientific study produces evidence that a chemical causes cancer in animals, can we can then be reasonably sure that the chemical will cause cancer in humans?

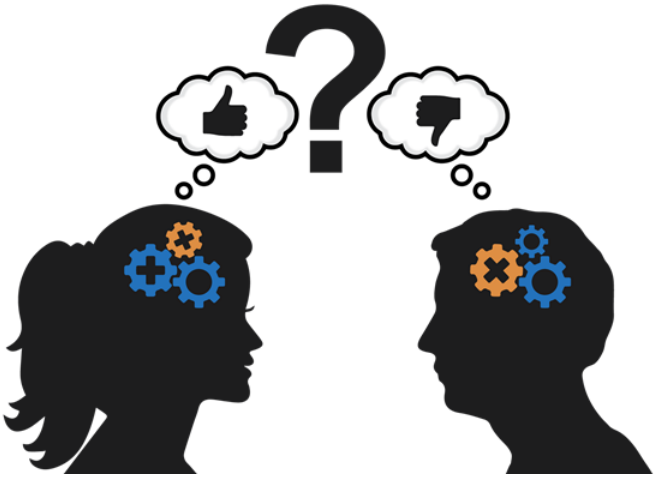


Figure 2.6.6.15.1.2. Understanding and application of toxicology sometimes involves elements of expert judgment and intuition (Image Source: Adapted from iStock Photos, ©)

Examples of Findings from Intuitive Toxicology

Examples of the findings from studies of intuitive toxicology in the United States, Canada, and the United Kingdom include:

- The public is more likely than toxicologists to think chemicals pose greater risks.
- The public finds it difficult to understand the concept of dose–response relationships.
- Much disagreement between toxicologists about how to interpret various results.
- Technical judgments of toxicologists were also found to be associated with factors such as their type of employment (for example, academia, government, or industry), gender, and age.

These types of studies have identified misconceptions that experts should try to clarify in interactions with the public. The results also suggest that disagreement among experts, especially as perceived by the news media and the public, can play a key role in controversies over toxicology-related risks.

Learn more about Intuitive Toxicology

- Perception of risk.
  Slovic P.
  Science. 1987 Apr 17;236(4799):280-5.
- Risk perception: it's personal.
  Brown VJ. Environ Health Perspect. 2014 Oct;122(10):A276-9. doi: 10.1289/ehp.122-A276
- Intuitive toxicology: expert and lay judgments of chemical risks. Neil N, Malmfors T, Slovic P. Toxicol Pathol. 1994 Mar-Apr;22(2):198-201.
- Judgments of chemical risks: comparisons among senior managers, toxicologists, and the public. Mertz CK, Slovic P, Purchase IF. Risk Anal. 1998 Aug;18(4):391-404.

Knowledge Check

1) Intuitive toxicology studies show that:

a) All toxicologists think the same
b) Members of the public usually think like toxicologists
c) There can be meaningful differences among toxicologists in how they look at the same set of toxicology study results
d) Intuitive toxicology is not important to consider in communication efforts

Answer

There can be meaningful differences among toxicologists in how they look at the same set of toxicology study results - **This is the correct answer.**

Intuitive toxicology studies show that there can be meaningful differences among toxicologists in how they look at the same set of toxicology study results.

2) Which of the following statements is correct?

a) The concept of dose-response relationships is easily understood by the public

b) The public and toxicologists tend to agree about the risks of chemicals

c) Technical judgments of toxicologists have been found to not be associated with factors such as their type of employment (for example, academia, government of industry, gender, and age)

d) Technical judgments of toxicologists have been found to be associated with factors such as their type of employment (for example, academia, government of industry, gender, and age)

Answer

Technical judgments of toxicologists have been found to be associated with factors such as their type of employment (for example, academia, government of industry, gender, and age) - **This is the correct answer.**

Technical judgments of toxicologists have been found to be associated with factors such as their type of employment (for example, academia, government of industry, gender, and age).

This page titled 2.6.6.15.1: Intuitive Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **15.1: Intuitive Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.15.2: Risk Communication

Risk Communication

Risk communication is the exchange of information about risks.

Rules for Communicating Risk

Much information about how risks could be communicated is available. Some key points about risk communication are identified in the "Seven Cardinal Rules for Communicating Risk" from the work of Dr. Vincent Covello and used by U.S. EPA and others:

- Accept and involve the public as a legitimate partner.
- Listen to the public's specific concerns.
- Be honest, frank, and open.
- Coordinate and collaborate with other credible sources.
- Meet the needs of the media.
- Speak clearly and with compassion.
- Plan carefully and evaluate your efforts.

Lessons Learned About Communicating Risk

Some of the lessons that organizations have learned about communicating exposure and health effects information to study subjects, the community, and the public include:

- Communication is not a "cheap add-on" to a study. It must be planned and budgeted at the start. The researcher must know the community and establish relationships early in the project. Communications should be tailored to the project and should contain what people really need to know. The study results that are most significant for the community should be emphasized. Moreover, results should be communicated in a format and a manner that subjects can readily understand. Researchers should evaluate and learn from each study.
- Ignoring communication may lead to legal problems.
- Communicating risk is part of societal accountability.
- Principles and guidelines, including proper terminology, are needed.
- Guidelines should be enforceable.
- Communication requires resources.
- It should be determined early in the project who has control of the release of results, and whether results will be presented in stages or all at once.
- A professional's credibility is at risk when decisions about communication of study results are being made.
- Mechanisms may be needed to proactively consider communication.
- The role of Institutional Review Boards (IRBs) must be considered in developing communication.

Learn more about communicating risk

Lessons Learned from a Crisis and Emergency

Six principles of effective crisis and risk communication are:

1. Be first
2. Be right
3. Be credible
4. Express empathy
5. Promote action
6. Show respect

"The CDC acknowledges that less-than-clear communication about what was known and not known about the possible health effects of the Elk River spill may have affected communities' trust in government." Learn more



Figure 2.6.6.15.2.1. Charleston, West Virginia viewed from across the Kanawha River, of which the Elk River is a tributary
(Image Source: iStock Photos, ©)

Uncertainty

Uncertainty is defined as "imperfect knowledge concerning the present or future state of an organism, system, or (sub)population under consideration." In other sources (EFSA, 2018), "uncertainty is defined as referring to all types of limitations in the knowledge available to assessors at the time an assessment is conducted and within the time and resources available for the assessment." There are different types of uncertainty, some quantifiable and others not, some reducible and others not.

Due to lack of knowledge, variability adds to the overall uncertainty. Ignoring uncertainty may lead to incomplete risk assessments, poor decision-making, and poor risk communication (European Commission, 2015). The degree to which characterization of uncertainty (and variability) is needed will depend on the risk assessment and risk management contexts as determined in the questions asked (problem formulation).



Figure 2.6.6.15.2.1. Uncertainty
(Image Source: Adapted from iStock Photos, ©)

Uncertainty should be acknowledged and described, such as outlining any data gaps or issues relating to methodology. What is being done to address the areas of uncertainty is also important. In its guideline When Food Is Cooking Up a Storm, the European Food Safety Authority provides a framework to assist decision-making about appropriate communications approaches in a wide variety of situations that can occur when assessing and communicating on risks related to food safety in Europe. It is directed towards governmental agencies that regulate the food sector.

EFSA has developed a harmonized approach to assessing and taking account of uncertainties in food safety, and animal and plant health. In 2018, EFSA published its Guidance on Uncertainty Analysis in Scientific Assessment which offers a diverse toolbox of scientific methods and technical tools for uncertainty analysis. It is sufficiently flexible to be implemented in such diverse areas as plant pests, microbiological hazards and chemical substances. Further, in a separate document EFSA (2018) describes the principles and methods behind its guidance. It provides a flexible framework within which different methods may be selected, according to the needs of each risk assessment. It is recommended that assessors should systematically identify sources of uncertainty, checking each part of their assessment to minimize the risk of overlooking important uncertainties.

Communicating Uncertainty in Risk Assessments and in Risk Management



Figure 2.6.6.15.2.3. Uncertainty
(Image Source: Adapted from iStock Photos, ©)

By late 2018, EFSA is expected to have a practical guidance for communication specialists on how to communicate the results of uncertainty analysis to different target audiences, including the public. The document aims to help EFSA to communicate scientific uncertainties to its different audiences by using more accessible language tailored to their needs.

Learn more about uncertainty and communicating about it

- U.S. FDA: "Elemental Analysis Manual for Food and Related Products" - https://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm2006954.htm
- U.S. FDA: "Best Practices in Risk Communication and Communicating Uncertainty: Applications to FDA-Regulated Products" - https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/RiskCommunicationAdvisoryCommittee/UCM486965.pdf
- EFSA: "Risk, Precautionary Principle and Uncertainty. European Commission Uncertainty assessment"
- EFSA: Workshop on the trial of the EFSA guidance document on uncertainty analysis in scientific assessments, 22 - 23 June 2017, Parma, Italy. https://www.efsa.europa.eu/en/supporting/pub/en-1313
- EFSA: Member State multilingual online survey on communicating uncertainty to different target audiences. European Food Safety Authority - http://www.efsa.europa.eu/en/supporting/pub/en-1413
- EFSA: "Communicating uncertainty: some applications of indicators and their validity" [PDF]
- EFSA: "When Food Is Cooking Up a Storm – Proven Recipes for Risk Communications"

Knowledge Check

1) Which of the following is not true about communicating risk to a community about exposures and health effects?

a) Results should be communicated in a format and a manner that subjects can readily understand
b) Results do not need to be communicated in a format and a manner that subjects can readily understand
c) Communications should be tailored to the project and should contain what people really need to know
d) Researchers and others can learn from studying good and bad risk communication efforts

Answer

Results do not need to be communicated in a format and a manner that subjects can readily understand - **This is the correct answer.**

2) According to the European Commission, ignoring uncertainty may lead to:

a) Great decision-making
b) Poor decision-making
c) Effective risk communication
d) Use of the most accurate knowledge available

Answer

Poor decision-making - **This is the correct answer.**
According to the European Commission, ignoring uncertainty may lead to poor decision-making.

This page titled 2.6.6.15.2: Risk Communication is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **15.2: Risk Communication** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

SECTION OVERVIEW

2.6.6.16: Environmental Toxicology, Environmental Health, and One Health

Learning Objectives

After completing this lesson, you will be able to:

- Define Environmental Toxicology.
- Describe the differences between environmental toxicology and environmental health.
- Define One Health.

In this section...

Topics include:

2.6.6.16.1: Environmental Toxicology

2.6.6.16.2: Environmental Health

2.6.6.16.3: One Health

Section 16: Key Points

What We've Covered

This section made the following main points:

- Environmental Toxicology:
  - Is the multidisciplinary study of the effects of manmade and natural chemicals on health and the environment.
  - Includes work in academia, companies, government agencies, and elsewhere.
- Environmental Health:
  - Is a branch of public health focused on the relationships between people and their environment.
  - Seeks to advance policies and programs to reduce chemical and other environmental exposures in air, water, soil, and food.
  - Saves lives, saves money, and protects your future.
- One Health:
  - Is a global concept and strategy recognizing the links between the health of people, animals, and the environment.
  - Examines changes in the interactions between people, animals, and the environment and how these changes impact human and animal health.

This page titled 2.6.6.16: Environmental Toxicology, Environmental Health, and One Health is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

2.6.6.16.1: Environmental Toxicology

Environmental Toxicology

Environmental Toxicology is the multidisciplinary study of the effects of manmade and natural chemicals on health and the environment. This includes the study of the effects of chemicals on organisms in their natural environments and in the ecosystems to which they belong.

Branches of Environmental Toxicology

Environmental Toxicology covers a wide range of interdisciplinary studies, as illustrated in Figure 1:

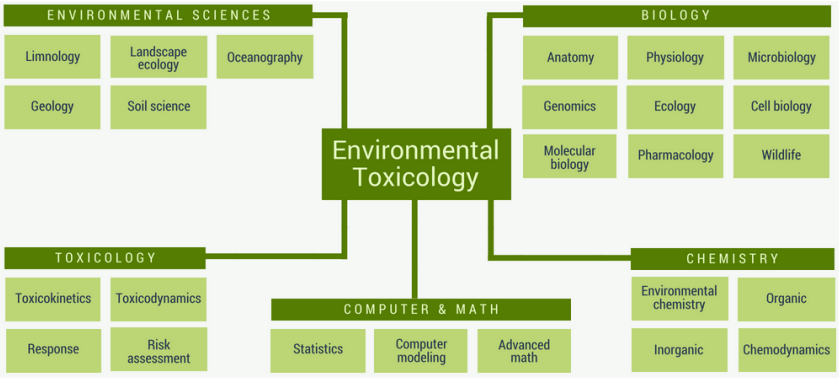


Figure 2.6.6.16.1.1. Environmental Toxicology interdisciplinary core (not comprehensive)
(Image Source: Adapted from Wikipedia under the Creative Commons Attribution-ShareAlike 3.0 License)

Scope of Work and Study

Environmental toxicologists work in academia, companies, government agencies, and elsewhere. The work can include laboratory studies, computer modeling, and work "in the field." It is not unusual for an environmental toxicologist to also have training in other areas—for example, public health, environmental chemistry, and pharmacology.
Some examples of what Environmental Toxicologists study include:

- The effects of a chemical or other substance at various concentrations on various species.
- Whether a chemical or other substance can bioaccumulate (increase over time) in animals or other organisms. This is important for human exposures if the bioaccumulation occurs in animals that are part of the human food chain, such as fish.
- Emerging issues such as the study of the sources and effects of microplastics that could become part of the human food chain.

Learn more about microplastics

- U.S. EPA "Toxicological Threats of Plastic" - https://www.epa.gov/trash-free-waters/toxicological-threats-plastic
- GreenFacts "Marine Litter" - https://www.greenfacts.org/en/marine-litter/index.htm#1
- European Chemicals Agency "Microplastics" - https://echa.europa.eu/hot-topics/microplastics



Figure 2.6.6.16.1.2. Microplastics are plastic debris less than five millimeters in length
(Image Source: National Ocean Service, National Oceanic and Atmospheric Administration. Original image)

Another emerging global issue is the health of bees. In the news in recent years are terms like the Colony Collapse Disorder (CCD), pesticides like the neonicotinic pesticides (also called neonicotinoids), and parasites that only reproduce in bee colonies. About 75% of all flowering plants rely on animal pollinators and about one-third of our food production is dependent on animal pollinators. The general declining health of honeybees and other bees is thought to be related to complex interactions among multiple stressors, including pesticides and parasites, and stressors like poor nutrition due to declining foraging habitats, bee management practices, and a lack of genetic diversity.

Neonicotinoids have been a focus of international attention and their mode of action is on the central nervous system of insects. Neonicotinoids are highly toxic to honeybees and also native bees like bumble bees and blue orchard bees, and sub-lethal levels can affect foraging and the ability to reproduce. Further, neonicotinoids can be persistent in the environment, and can be absorbed by plants and found in pollen and nectar.



Figure 2.6.6.16.1.3. A honeybee gathers pollen from a flower
(Image Source: iStock Photos, ©)

Learn more about pollinators

- U.S. Fish and Wildlife Service "Pollinators" - https://www.fws.gov/Pollinators/
- U.S. Department of Agriculture, Forest Service "Pollinators" - https://www.fs.fed.us/wildflowers/pollinators/index.shtml
- U.S. EPA "Pollinator Health Concerns" - https://www.epa.gov/pollinator-protection/pollinator-health-concerns
- European Commission "The EU Approach to Tackle Pollinator Decline" - http://ec.europa.eu/environment/nature/conservation/species/pollinators/index_en.htm
- European Food Safety Authority "Bees Under Siege: Making Sense of Multiple Stressors" - https://www.youtube.com/watch?v=ZVKJNc0tBDM&feature=youtu.be&list=PLGDvgn1aAEEbWUxOzO8FjrtBQpLiC5DOB
- Xerces Society for Invertebrate Conservation "Neonicotinoids and Bees" - http://xerces.org/neonicotinoids-and-bees/

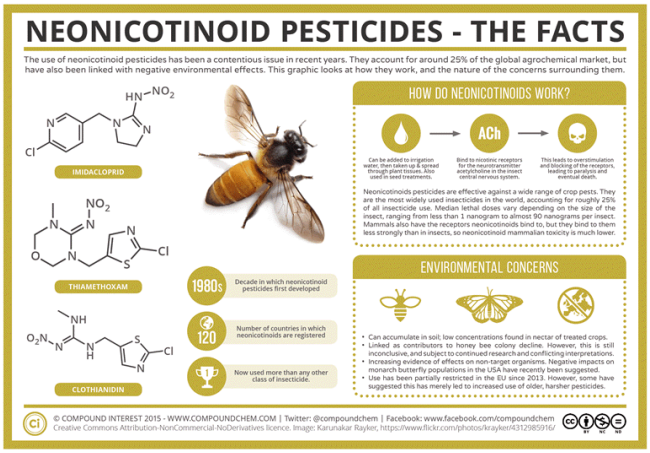


Figure 2.6.6.16.1.4. Neonicotinoid Pesticides
(Image Source: Compound Interest, © 2015 - used under Creative Commons Attribution-NonCommercial-NoDerivatives license. Bee image captured by Karunakar Rayker. Original Image. Original Infographic)

Knowledge Check

1) The interdisciplinary core ("branches") of environmental toxicology includes:

a) Environmental sciences; physics; toxicology; chemistry; biology
b) Environmental sciences; engineering; toxicology; biology; computer and math
c) Environmental sciences; computer and math; toxicology; biology; law
d) Environmental sciences; biology; toxicology; chemistry; computer and math

Answer

Environmental sciences; biology; toxicology; chemistry; computer and math - **This is the correct answer.**
The interdisciplinary core ("branches") of environmental toxicology includes environmental sciences; biology; toxicology; chemistry; computer and math.

2) Which issues in environmental toxicology relate to the human food supply?

a) Bioaccumulation of substances in fish
b) Effects of neonicotinic pesticides on bees
c) Both the bioaccumulation of substances in fish and effects of neonicotinic pesticides on bees

Answer

Both the bioaccumulation of substances in fish and effects of neonicotinic pesticides on bees - **This is the correct answer.**
Both the bioaccumulation of substances in fish and effects of neonicotinic pesticides on bees relate to the human food supply.

This page titled 2.6.6.16.1: Environmental Toxicology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **16.1: Environmental Toxicology** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.16.2: Environmental Health

Environmental Health

Environmental health is a branch of public health. It focuses on the relationships between people and their environment, and promotes human health and well-being. Further, it fosters healthy and safe communities and is a key part of any comprehensive public health system. The field works to advance policies and programs to reduce chemical and other environmental exposures in air, water, soil, and food to protect people and provide communities with an healthier environments.

Toxicology vs. Environmental Health

The two fields are closely connected, with a large intersection.
The terminology is loose. For example, is environmental health about the health of the environment and/or the effect of the environment on health?

Scope of Work and Study

Environmental health specialists identify and evaluate environmental hazards and their sources. They also limit exposures to physical, chemical, and biological agents in air, water, soil, food, and other environmental media or settings that may adversely affect human health (Source).



Figure 2.6.6.16.2.1. Environmental health saves lives, saves money, and protects your future
(Image Source: National Environmental Health Association (NEHA), © 2016. Original image)

National Center for Environmental Health (NCEH)

The CDC National Center for Environmental Health (NCEH) plans, directs, and coordinates a program to protect the American people from environmental hazards. The NCEH seeks to prevent premature death, avoidable illness, and disability caused by non-infectious, non-occupational environmental and related factors. One focus is on safeguarding the health of venerable populations, such as children, the elderly, and people with disabilities, from certain environmental hazards.

Learn more about environmental health

- U.S. CDC National Center for Environmental Health - **https://www.cdc.gov/nceh/**
- American Public Health Association (APHA): "Environmental Health" - **https://www.apha.org/topics-and-issues/environmental-health**
- National Environmental Health Association (NEHA): "Definitions of Environmental Health" - **https://www.neha.org/about-neha/definitions-environmental-health**

Knowledge Check

1) Which of the following are important environmental health issues?

a) Radon in air
b) Lead contamination in drinking water
c) Foodborne illness from contaminated food
d) All of these are important for human health

Answer

All of these are important for human health

This page titled 2.6.6.16.2: Environmental Health is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **16.2: Environmental Health** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.16.3: One Health

One Health

One Health is a worldwide concept and strategy recognizing that the health of people, animals, and the environment are all connected. For example, the Centers for Disease Control and Prevention (CDC) works with physicians, veterinarians, ecologists, and many others to monitor and control public health threats and to learn about how diseases spread among people, animals, and the environment.

Link Between Human, Animal, and Environmental Health

One Health is important at the local, regional, national, and global levels, and there are many examples of its importance. One example of how human, animal, and environmental health are linked involves bacteria, cows, farms, food, lettuce, and humans:



Cows graze next to a lettuce field. Cows can carry *E. coli* but still look healthy.



E. coli from cow manure in the nearby farm can contaminate the lettuce field.



People eat contaminated lettuce and can become infected with *E. coli*. Serious illness or sometimes death can result.

Figure 2.6.6.16.3.1. Human, animal, and environmental health are linked
(Image Source: CDC - original image)

Another example of One Health involving animals and humans is the shared susceptibility to some diseases and environmental hazards. Animals can serve as early warning signs of potential human illness. An example is birds dying from West Nile virus before people in the same areas get sick from exposures to this virus.

Factors that Affect Human and Animal Health

Some interactions between people, animals, and the environment have changed in recent years and these changes have impacted animal and human health.

Factor (Cause)	Change (Effect)
Human populations are growing and expanding into new geographic areas.	More people live in close contact with wild and domestic animals. Close contact provides more opportunities for diseases to pass between animals and people.
The earth has experienced changes in climate and land use, such as deforestation and intensive farming practices.	Disruptions in environmental conditions and habitats provide new opportunities for diseases to pass to animals.
International travel and trade have increased.	Diseases can spread quickly across the globe.

Table 2.6.6.16.3.1. Factors that affect human and animal health
(Source: CDC: One Health Basics)

Learn more about One Health

- U.S. CDC: "One Health: From Concept to Action" - https://youtu.be/TG0pduAYESA
- U.S. CDC: "One Health Basics" - https://www.cdc.gov/onehealth/basics/index.html
- World Organization for Animal Health (OIE): "One Health: By Protecting Animals, We Preserve Our Future [PDF] - http://www.oie.int/fileadmin/Home/eng/Media_Center/img/Infographies/A4-EN-WEB.pdf
- U.S. FDA: "One Health: It's for All of Us" - https://www.fda.gov/AnimalVeterinary/ResourcesforYou/AnimalHealthLiteracy/ucm278139.htm

Knowledge Check

1) One Health is a concept and strategy recognizing that the health of _____, _____, and _____ are all connected.

a) People, plants, the environment
b) People, animals, the environment
c) People, animals, microbes

Answer

People, animals, the environment - **This is the correct answer.**
One Health is a concept and strategy recognizing that the health of people, animals, and the environment are all connected.

2) Which of the following factors that affect human and animal health is not correct?

a) Fewer people in recent years live in close contact with wild and domestic animals
b) Disruptions in environmental conditions and habitats provide new opportunities for diseases to pass to animals
c) International travel and trade have increased, and diseases can spread quickly across the globe

Answer

Fewer people in recent years live in close contact with wild and domestic animals - **This is the correct answer.**
The statement "Fewer people in recent years live in close contact with wild and domestic animals" is incorrect.

This page titled 2.6.6.16.3: One Health is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **16.3: One Health** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

2.6.6.17: Conclusion

Thank you for completing ToxTutor! We trust it has given you a strong foundation of understanding in this important area of science.

In conclusion, we have now covered the following topics throughout this course:

1. Introduction to Toxicology
2. Dose and Dose Response
3. Toxic Effects
4. Interactions
5. Toxicity Testing Methods
6. Risk Assessment
7. Exposure Standards and Guidelines
8. Basic Physiology
9. Introduction to Toxicokinetics
10. Absorption
11. Distribution
12. Biotransformation
13. Excretion
14. Cellular Toxicology
15. Intuitive Toxicology and Risk Communication
16. Environmental Toxicology, Environmental Health, and One Health

Contact Information

If you would like additional information about toxicology and environmental health, or if you have any questions or comments, please contact us at **toxmsdt@gmail.com**

Credits

ToxTutor was adopted from the **U.S National Library of Medicine** in 2021. More Information can be seen in the ToxTutor **Bibliography**.



This page titled 2.6.6.17: Conclusion is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by ToxMSDT Online component via source content that was edited to the style and standards of the LibreTexts platform.

- **Section 17: Conclusion** by ToxMSDT Online component is licensed CC BY-NC 4.0. Original source: http://www.toxmsdt.com.

CHAPTER OVERVIEW

Back Matter

Detailed Licensing

Detailed Licensing

CHAPTER OVERVIEW

Back Matter

Index
Glossary
Detailed Licensing

Index

A

acid rain
- 2.2.1: Acid Rain
- 2.2.1.2: Acid Rain Transport
- 2.2.1.3: Acid Snow
- 7.2.2.2: Acid Deposition

acid snow
- 2.2.1.3: Acid Snow

adaptive radiation
- 4.2.1: Extinction

Aeration
- 2.2.2.10: Water Treatment

age structure
- 5.1: The Human Population

agriculture
- 2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture

amphiprotic specie
- 2.2.2.8: Water Chemistry

anthrosphere
- 2.5.8.6: The Anthrosphere

Aposematic Coloration
- 3.3.1.2: Competition

aquatic biomes
- 3.5.3: Aquatic Biomes

Aquatic Chemistry
- 2.2.2: Aquatic Chemistry

aquifer
- 2.5.8.2: The Hydrosphere

atmosphere
- 2.2.3.1: Atmosphere

B

basic
- 2.2.4: Biochemical Cycles
- 2.2.4.2: Carbon Cycle
- 2.2.4.3: Nitrogen Cycle

Batesian mimicry
- 3.3.1.2: Competition

biochar
- 2.5.11.11: Soil Loss - Desertification and Deforestation

biochemical effect
- 2.5.16.4: Toxic Substances and Toxicology

biodegradation
- 2.5.12.10: Biodegradation

biodiversity
- 4.2.1: Extinction

biodiversity hotspot
- 4.2.1: Extinction

Biogeochemical Evolution
- 2.1.4.2: Biogeochemical Evolution

biological agents
- 5.4.1: Types of Environmental Hazards

biome
- 3.5.1: Climate and Biomes

biosphere
- 2.1.4: The Biosphere
- 2.5.8.5: The Biosphere
- 2.5.12.2: The Biosphere

biotransformation
- 2.6.6.12.1: Introduction to Biotransformation

bond energy
- 2.5.4.5: Covalent Bonds in Compounds

botulinum
- 2.5.16.6: Protecting Water, Food, and Air

breeder reactor
- 2.5.15.10: Nuclear Energy

C

camouflage
- 3.3.1.2: Competition

carbon cycle
- 2.2.3.2: Carbon Cycle
- 2.2.4.2: Carbon Cycle

carbon sequestration
- 2.5.10.5: The Enormous Importance of Climate

carrying capacity
- 3.2.3: Population Growth and Regulation

cellulose
- 2.5.14.7: Cellulose

Centers for Disease Control and Prevention
- 5.4.2: Epidemiology

CFCs
- 2.2.3.4: Ozone

chain reaction
- 2.5.15.10: Nuclear Energy

chemical diversity
- 4.2.1: Extinction

climate
- 2.5.10.5: The Enormous Importance of Climate

climate change
- 2.5.10.5: The Enormous Importance of Climate

climax community
- 3.3.1.2: Competition

clod
- 2.5.11.8: Have You Thanked a Clod Today?

coagulation
- 2.2.2.10: Water Treatment

cometabolism
- 2.5.12.10: Biodegradation

commensalism
- 3.3.1.2: Competition
- 3.3.1.4: Symbiosis

community
- 3.3.1: Biotic Interactions

community ecology
- 3.3.1: Biotic Interactions

competitive exclusion principle
- 3.3.1: Biotic Interactions
- 3.3.1.2: Competition

conservation tillage
- 2.5.11.11: Soil Loss - Desertification and Deforestation
- 2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture

Crop rotations
- 5.3.4: Sustainable Agriculture

Curtius rearrangement
- 2.4.3.4: Green Technologies for Safer Chemical Production

cyclodextrins
- 2.4.5.3: Specific Replacements

cytotoxicity assays
- 2.6.3.4: Different Cytotoxicity Assays

D

DDT
- 2.4.1.2: Green Chemistry Concepts
- 5.4.1: Types of Environmental Hazards

deforestation
- 2.5.11.11: Soil Loss - Desertification and Deforestation

desertification
- 2.5.11.11: Soil Loss - Desertification and Deforestation

detoxication
- 2.5.12.10: Biodegradation

DNA replication
- 2.6.2.3: DNA and RNA Metabolism

Dobson Unit (DU)
- 2.2.3.4: Ozone

E

ecological footprint
- 8.2: History of Sustainability

ecosystem diversity
- 4.2.1: Extinction

Emsleyan/Mertensian mimicry
- 3.3.1.2: Competition

Endemic diseases
- 5.4.2: Epidemiology

Endemic species
- 4.2.1: Extinction

Environmental Chemistry
- 2: Environmental Chemistry

environmental disturbance
- 3.3.1.2: Competition

Environmental Toxicology
- 5.4.4: Environmental Toxicology

Epidemic diseases
- 5.4.2: Epidemiology

epidemiology
- 5.4.2: Epidemiology

eutrophication
- 2.2.4.1: Biochemical Cycles

Exemplar
- 2.2.1.4: Electricity Generation
- 2.2.1.6: Sources of Nitrogen Oxides
- 2.2.1.7: Sources of Sulfur Oxides

exponential growth
- 3.2.3: Population Growth and Regulation

extinction
- 4.2.1: Extinction

extinction rate
- 4.2.1: Extinction

F

Flucculation
- 2.2.2.10: Water Treatment

food security
- 5.3.1: Food Security

foundation species
- 3.3.1.2: Competition

frontier ethic
8.3.2: Perspectives of the Environment

G

Gaia hypothesis
2.1.4.3: Gaia - Bioregulation of the Environment
genetic diversity
4.2.1: Extinction
genetic engineering
5.3.3: Selective Breeding and Genetic Engineering
geochemistry
2.1: Geochemistry (Lower)
geosphere
2.5.8.4: The Geosphere
2.5.11.2: The Nature of the Geosphere
glyphosate
2.5.12.1: Pigweed's Revenge
green chemistry
2.4: Key Elements of Green Chemistry (Lucia)
green products
2.5.1.10: Green Products and Services- Design for Sustainability

H

Herbivory
3.3.1: Biotic Interactions
heterogeneity
4.2.1: Extinction
host
3.3.1.2: Competition
human population
5.1: The Human Population
humification
2.5.11.8: Have You Thanked a Clod Today?
hydrogen
2.5.3.2: Hydrogen - The Simplest Atom
hydrologic cycle
2.5.8.2: The Hydrosphere
Hydrosphere
2.2.2.6: Natural Water

I

incidence
5.4.2: Epidemiology
integrated pest management
5.3.4: Sustainable Agriculture
intercropping
5.3.4: Sustainable Agriculture
island biogeography
3.3.1.2: Competition

K

keystone species
3.3.1.2: Competition
Koch's postulates
5.4.2: Epidemiology

L

Lattice Energy
2.5.4.3: Sodium Chloride and Ionic Bonds
life cycles
2.4.2.1: Introduction to Life Cycles
lignin
2.5.14.9: Lignin
logistic growth
3.2.3: Population Growth and Regulation
lysergic acid
2.4.1.2: Green Chemistry Concepts

M

Müllerian Mimicry
3.3.1.2: Competition
megafauna
4.2.1: Extinction
Mertensian mimicry
3.3.1.2: Competition
microcrystalline cellulose
2.5.14.7: Cellulose
micronutrients
2.5.11.9: Plant Nutrients and Fertilizers in Soil
mimicry
3.3.1: Biotic Interactions
mineralization
2.5.12.10: Biodegradation
minimal tillage
5.3.4: Sustainable Agriculture
moderator
2.5.15.10: Nuclear Energy
morbidity
5.4.2: Epidemiology
morbidity rates
5.4.2: Epidemiology
mortality
5.4.2: Epidemiology
mortality rates
5.4.2: Epidemiology
municipal refuse
2.5.11.7: The Geosphere as a Waste Repository
mutualism
3.3.1.2: Competition
3.3.1.4: Symbiosis

N

nuclear waste
2.5.15.10: Nuclear Energy
nucleosynthesis
2.1.1.2: Origin of the Elements

O

octet
2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons
octet rule
2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons
organic farming
5.3.4: Sustainable Agriculture
overpopulation
5.1.3: Looking Ahead
ozone
2.2.3.4: Ozone
ozone depletion
7.2.2.1: Ozone Depletion

P

paleosols
3.4.4: Soils
pandemic diseases
5.4.2: Epidemiology
parasite
3.3.1.2: Competition
parasitism
3.3.1.4: Symbiosis
passive stability
2.5.15.10: Nuclear Energy
Permanent Hard Water
2.2.2.6: Natural Water
photosynthesis
2.2.3.5: Photosynthesis
pic
2.5.7.11: Questions and Problems
pioneer species
3.3.1.2: Competition
population growth
3.2.3: Population Growth and Regulation
Prebiotic Atmosphere
2.1.3.2: Origin and Evolution of the Atmosphere
precautionary principle
8.2: History of Sustainability
predator
3.3.1: Biotic Interactions
prevalence
5.4.2: Epidemiology
primary succession
3.3.1.2: Competition
pseudogenes
3.3.1.4: Symbiosis

Q

QSAR
2.6.4.4: Other Approaches for Predictive Toxicity Modeling
Quantitative Structure Activity Relationships
2.6.4.4: Other Approaches for Predictive Toxicity Modeling

R

relative species abundance
3.3.1.2: Competition
Remediation
7.2.1.1: Water Pollutants and Their Sources
rhizosphere
2.5.11.8: Have You Thanked a Clod Today?
ricin
2.5.16.6: Protecting Water, Food, and Air

S

sanitary landfills
2.5.11.7: The Geosphere as a Waste Repository
sarin
2.5.16.6: Protecting Water, Food, and Air
secondary succession
3.3.1.2: Competition
sedimentation
2.2.2.10: Water Treatment

sevin
2.4.3.4: Green Technologies for Safer Chemical Production
soil conservation
2.5.11.11: Soil Loss - Desertification and Deforestation
soil erosion
2.5.11.11: Soil Loss - Desertification and Deforestation
Soil processes
3.4.4: Soils
Soil Profiles
3.4.4: Soils
species richness
3.3.1.2: Competition
Sporadic diseases
5.4.2: Epidemiology
sulfur oxides
2.2.1.7: Sources of Sulfur Oxides
surface albedo
2.5.11.2: The Nature of the Geosphere
sustainability
2.5.1.3: Seeing Green
sustainable agriculture
5.3.4: Sustainable Agriculture
symbiosis
3.3.1: Biotic Interactions
3.3.1.2: Competition
3.3.1.4: Symbiosis
T
Temporary Hard Water
2.2.2.6: Natural Water
terrestrial biomes
3.5.2: Terrestrial Biomes
Total Fertility Rate
5.1.3: Looking Ahead
transmutation
2.5.15.10: Nuclear Energy
travelling wave reactor
2.5.15.10: Nuclear Energy
W
water
2.2.2: Aquatic Chemistry
2.2.2.5: Fundamental Characteristics of Water
2.2.2.6: Natural Water
2.2.2.8: Water Chemistry
X
xenobiotic substances
2.5.16.4: Toxic Substances and Toxicology

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: ENVS C100: Environmental Science (Hoerer)

Webpages: 679

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC-SA 4.0: 38.7% (263 pages)
- CC BY-NC 4.0: 29.2% (198 pages)
- CC BY 4.0: 12.5% (85 pages)
- Undeclared: 11.5% (78 pages)
- CC BY-SA 4.0: 5.4% (37 pages)
- CC BY-NC 3.0: 2.7% (18 pages)

By Page

- ENVS C100: Environmental Science (Hoerer) - *Undeclared*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Introduction - *CC BY-NC 4.0*
    - 1.1: Environmental Science Preface - *Undeclared*
    - 1.2: What Is Science, and How Does It Work? - *CC BY-NC-SA 4.0*
      - 1.2.1: What is Science? - *CC BY-NC-SA 4.0*
      - 1.2.2: The Scientific Method - *CC BY-SA 4.0*
      - 1.2.3: Scientific Papers - *CC BY-NC 4.0*
      - 1.2.4: Basic and Applied Science - *CC BY-SA 4.0*
      - 1.2.5: Data Dive- Tonsillectomy Trends - *Undeclared*
      - 1.2.6: Review - *CC BY-NC 4.0*
    - 1.3: Math Blast- An Overview of Essential Mathematics Used in Science - *CC BY-NC 4.0*
      - 1.3.1: Getting down the basics- What are the essential mathematical and statistical terms and concepts? - *Undeclared*
      - 1.3.2: Summarizing the data- Descriptive statistics - *CC BY-NC 4.0*
      - 1.3.3: Putting it all together- Inferential statistics and hypothesis testing - *CC BY-NC 4.0*
      - 1.3.4: Data Dive- California Wildfires - *Undeclared*
      - 1.3.5: Review - *CC BY-NC 4.0*
  - 2: Environmental Chemistry - *CC BY-NC-SA 4.0*
- 00: Front Matter - *Undeclared*
  - TitlePage - *Undeclared*
  - InfoPage - *Undeclared*
  - Table of Contents - *Undeclared*
  - Licensing - *Undeclared*
- 2.1: Geochemistry (Lower) - *CC BY-NC 3.0*
  - Front Matter - *Undeclared*
    - Licensing - *Undeclared*
  - 2.1.1: The Earth and its Lithosphere - *CC BY-NC 3.0*
    - 2.1.1.1: Composition and Structure of the Earth - *CC BY-NC 3.0*
    - 2.1.1.2: Origin of the Elements - *CC BY-NC 3.0*
    - 2.1.1.3: Formation and evolution of the Earth - *CC BY-NC 3.0*
    - 2.1.1.4: The Earth's crust - *CC BY-NC 3.0*
  - 2.1.2: The Hydrosphere - *CC BY-NC 3.0*
    - 2.1.2.1: Water, Water Everywhere... - *CC BY-NC 3.0*
    - 2.1.2.2: The hydrosphere and the oceans - *CC BY-NC 3.0*
    - 2.1.2.3: Chemistry and geochemistry of the oceans - *CC BY-NC 3.0*
    - 2.1.2.4: Chemical budgets of oceanic elements - *CC BY-NC 3.0*
  - 2.1.3: The Atmosphere - *CC BY-NC 3.0*
    - 2.1.3.1: Structure and Composition of the Atmosphere - *CC BY-NC 3.0*

- 2.1.3.2: Origin and Evolution of the Atmosphere - *CC BY-NC 3.0*
- 2.1.4: The Biosphere - *CC BY-NC 3.0*
  - 2.1.4.1: Chemistry and Energetics of the Life Process - *CC BY-NC 3.0*
  - 2.1.4.2: Biogeochemical Evolution - *CC BY-NC 3.0*
  - 2.1.4.3: Gaia - Bioregulation of the Environment - *CC BY-NC 3.0*
- 2.2: Supplemental Modules (Environmental Chemistry) - *CC BY-NC-SA 4.0*
  - 2.2.1: Acid Rain - *CC BY-NC-SA 4.0*
    - 2.2.1.1: Acid Rain - *CC BY-NC-SA 4.0*
    - 2.2.1.2: Acid Rain Transport - *CC BY-NC-SA 4.0*
    - 2.2.1.3: Acid Snow - *CC BY-NC-SA 4.0*
    - 2.2.1.4: Electricity Generation - *CC BY-NC-SA 4.0*
    - 2.2.1.5: Playground 1 - *Undeclared*
    - 2.2.1.6: Sources of Nitrogen Oxides - *CC BY-NC-SA 4.0*
    - 2.2.1.7: Sources of Sulfur Oxides - *CC BY-NC-SA 4.0*
  - 2.2.2: Aquatic Chemistry - *CC BY-NC-SA 4.0*
    - 2.2.2.1: 1. Acid-Base Chemistry of Natural Aquatic Systems - *CC BY-NC-SA 4.0*
    - 2.2.2.2: 2. Carbonate Equilibria in Natural Waters - *CC BY-NC-SA 4.0*
    - 2.2.2.3: 3. Redox Equilibria in Natural Waters - *CC BY-NC-SA 4.0*
    - 2.2.2.4: 4. Solids in Contact With Natural Waters - *CC BY-NC-SA 4.0*
    - 2.2.2.5: Fundamental Characteristics of Water - *CC BY-NC-SA 4.0*
    - 2.2.2.6: Natural Water - *CC BY-NC-SA 4.0*
    - 2.2.2.7: Water Biology - *CC BY-NC-SA 4.0*
    - 2.2.2.8: Water Chemistry - *CC BY-NC-SA 4.0*
    - 2.2.2.9: Water Physics - *CC BY-NC-SA 4.0*
    - 2.2.2.10: Water Treatment - *CC BY-NC-SA 4.0*
  - 2.2.3: Atmospheric Chemistry - *CC BY-NC-SA 4.0*
    - 2.2.3.1: Atmosphere - *CC BY-NC-SA 4.0*
    - 2.2.3.2: Carbon Cycle - *CC BY-NC-SA 4.0*
    - 2.2.3.3: Carbon Oxides - *CC BY-NC-SA 4.0*
    - 2.2.3.4: Ozone - *CC BY-NC-SA 4.0*
    - 2.2.3.5: Photosynthesis - *CC BY-NC-SA 4.0*
  - 2.2.4: Biochemical Cycles - *CC BY-NC-SA 4.0*
    - 2.2.4.1: Biochemical Cycles - *CC BY-NC-SA 4.0*
    - 2.2.4.2: Carbon Cycle - *CC BY-NC-SA 4.0*
    - 2.2.4.3: Nitrogen Cycle - *CC BY-NC-SA 4.0*
- 2.3: Environmental Toxicology (van Gestel et al.) - *CC BY 4.0*
  - 00: Front Matter - *Undeclared*
    - Table of Contents - *Undeclared*
  - 2.3.1: Environmental Toxicology - *CC BY 4.0*
    - 2.3.1.1: Environmental toxicology - *CC BY 4.0*
    - 2.3.1.2: DPSIR - *CC BY 4.0*
    - 2.3.1.3: Short history - *CC BY 4.0*
  - 2.3.2: Environmental Chemistry, Chemicals - *CC BY 4.0*
    - 2.3.2.1: Introduction - *CC BY 4.0*
    - 2.3.2.2: Pollutants with specific properties - *CC BY 4.0*
    - 2.3.2.3: Pollutants with specific use - *CC BY 4.0*
  - 2.3.3: Environmental Chemistry - From Fate to Exposure - *CC BY 4.0*
    - 2.3.3.1: Environmental compartments - *CC BY 4.0*
    - 2.3.3.2: Sources of chemicals - *CC BY 4.0*
    - 2.3.3.3: Pathways and processes determining chemical fate - *CC BY 4.0*
    - 2.3.3.4: Partitioning and Partitioning Constants - *CC BY 4.0*
    - 2.3.3.5: Metal Speciation - *CC BY 4.0*
  - 2.3.4: Toxicology - *CC BY 4.0*
    - 2.3.4.1: Toxicokinetics - *CC BY 4.0*
    - 2.3.4.2: Toxicodynamics and Molecular Interactions - *CC BY 4.0*
    - 2.3.4.3: Toxicity Testing - *CC BY 4.0*
    - 2.3.4.4: Increasing ecological realism in toxicity testing - *CC BY 4.0*
    - 2.3.4.5: Availability and bioavailability - *CC BY 4.0*
    - 2.3.4.6: Degradation - *CC BY 4.0*
    - 2.3.4.7: Modelling exposure - *CC BY 4.0*
    - 2.3.4.8: Section 4.1. Toxicokinetics - *CC BY 4.0*
    - 2.3.4.9: Section 4.4. Increasing ecological realism in toxicity testing - *CC BY 4.0*
    - 2.3.4.10: Section 4.2. Toxicodynamics and Molecular Interactions - *CC BY 4.0*
    - 2.3.4.11: Section 4.3. Toxicity testing - *CC BY 4.0*
  - 2.3.5: Population, Community and Ecosystem Ecotoxicology - *CC BY 4.0*

- 2.3.5.1: Introduction- Linking population, community and ecosystem responses - *CC BY 4.0*
- 2.3.5.2: Population ecotoxicology in laboratory settings - *CC BY 4.0*
- 2.3.5.3: Wildlife population ecotoxicology - *CC BY 4.0*
- 2.3.5.4: Trait-based approaches - *CC BY 4.0*
- 2.3.5.5: Population models - *CC BY 4.0*
- 2.3.5.6: Metapopulations - *CC BY 4.0*
- 2.3.5.7: Community ecotoxicology - *CC BY 4.0*
- 2.3.5.8: Structure versus function incl. ecosystem services - *CC BY 4.0*
- 2.3.5.9: Landscape ecotoxicology - *CC BY 4.0*

- 2.3.6: Risk Assessment and Regulation - *CC BY 4.0*
  - 2.3.6.1: Introduction - The Essence of Risk Assessment - *CC BY 4.0*
  - 2.3.6.2: Ecosystem services and protection goals - *CC BY 4.0*
  - 2.3.6.3: Predictive risk assessment approaches and tools - *CC BY 4.0*
  - 2.3.6.4: Diagnostic risk assessment approaches and tools - *CC BY 4.0*
  - 2.3.6.5: Regulatory Frameworks - *CC BY 4.0*
  - 2.3.6.6: Risk management and risk communication - *CC BY 4.0*
  - 2.3.6.7: Risk perception - *CC BY 4.0*

- 2.4: Key Elements of Green Chemistry (Lucia) - *CC BY 4.0*
  - Front Matter - *CC BY-NC-SA 4.0*
    - Licensing - *Undeclared*
  - 2.4.1: Key Elements of Green Chemistry - *CC BY 4.0*
    - 2.4.1.1: Introduction to Green Chemistry - *CC BY 4.0*
    - 2.4.1.2: Green Chemistry Concepts - *CC BY 4.0*
    - 2.4.1.3: Case Study - *CC BY 4.0*
    - 2.4.1.4: Conclusions and Review - *CC BY 4.0*
  - 2.4.2: Life-Cycle Analysis - *CC BY 4.0*
    - 2.4.2.1: Introduction to Life Cycles - *CC BY 4.0*
    - 2.4.2.2: LCA/LCIA Concepts - *CC BY 4.0*
    - 2.4.2.3: Conclusions and Review - *CC BY 4.0*
  - 2.4.3: Hazards - *CC BY 4.0*
    - 2.4.3.1: Introduction to Hazards of Chemistry - *CC BY 4.0*
    - 2.4.3.2: Hazard Concepts - *CC BY 4.0*
    - 2.4.3.3: Case Study - Badger Army Ammunition Plant - *CC BY 4.0*
    - 2.4.3.4: Green Technologies for Safer Chemical Production - *CC BY 4.0*
    - 2.4.3.5: Conclusions and Review Questions - *CC BY 4.0*
  - 2.4.4: Alternative Solvents - *CC BY 4.0*
    - 2.4.4.1: Introduction to Solvents - *CC BY 4.0*
    - 2.4.4.2: Solvent Concepts - *CC BY 4.0*
    - 2.4.4.3: Solvent Alternatives - *CC BY 4.0*
    - 2.4.4.4: Conclusions and Review - *CC BY 4.0*
  - 2.4.5: Alternative Reagants - *CC BY 4.0*
    - 2.4.5.1: Introduction to Reagants - *CC BY 4.0*
    - 2.4.5.2: Reagant Concepts - *CC BY 4.0*
    - 2.4.5.3: Specific Replacements - *CC BY 4.0*
    - 2.4.5.4: Conclusions and Review - *CC BY 4.0*
  - 2.4.6: Reaction Types, Design, and Efficiency - *CC BY 4.0*
    - 2.4.6.1: Introduction to Reactions - *CC BY 4.0*
    - 2.4.6.2: Reaction Concepts - *CC BY 4.0*
    - 2.4.6.3: Reaction Design Concepts - *CC BY 4.0*
    - 2.4.6.4: Conclusions and Review - *CC BY 4.0*
  - Back Matter - *CC BY-NC-SA 4.0*
    - Detailed Licensing - *Undeclared*

- 2.5: Green Chemistry and the Ten Commandments of Sustainability (Manahan) - *CC BY-NC-SA 4.0*
  - Front Matter - *CC BY-NC-SA 4.0*
    - Licensing - *Undeclared*
  - 2.5.1: Sustainability and the Environment - *CC BY-NC-SA 4.0*
    - 2.5.1.1: Sustainability - *CC BY-NC-SA 4.0*
    - 2.5.1.2: The Environment and the Five Environmental Sphere - *CC BY-NC-SA 4.0*
    - 2.5.1.3: Seeing Green - *CC BY-NC-SA 4.0*
    - 2.5.1.4: Natural Capital of the Earth - *CC BY-NC-SA 4.0*
    - 2.5.1.5: Sustainability as a Group Effort- It Takes a (Very Big) Village - *CC BY-NC-SA 4.0*
    - 2.5.1.6: Sustainable Energy- Away from the Sun and Back Again - *CC BY-NC-SA 4.0*
    - 2.5.1.7: Green Science - *CC BY-NC-SA 4.0*
    - 2.5.1.8: Green Technology - *CC BY-NC-SA 4.0*
    - 2.5.1.9: Sustainability and the Eco-Economy - *CC BY-NC-SA 4.0*

- 2.5.1.10: Green Products and Services- Design for Sustainability - *CC BY-NC-SA 4.0*
- 2.5.1.11: Questions and Problems - *CC BY-NC-SA 4.0*

- 2.5.2: The Key Role of Chemistry and Making Chemistry Green - *CC BY-NC-SA 4.0*
  - 2.5.2.1: Chemistry is Good (and Unavoidable) - *CC BY-NC-SA 4.0*
  - 2.5.2.2: The Environment and its Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.2.3: What is Environmental Chemistry? - *CC BY-NC-SA 4.0*
  - 2.5.2.4: Environmental Pollution - *CC BY-NC-SA 4.0*
  - 2.5.2.5: Practice of Green Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.2.6: Green Chemistry and Synthetic Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.2.7: Reduction of Risk- Hazard and Exposure - *CC BY-NC-SA 4.0*
  - 2.5.2.8: The Risks of No Risks - *CC BY-NC-SA 4.0*
  - 2.5.2.9: Waste Prevention - *CC BY-NC-SA 4.0*
  - 2.5.2.10: Basic Principles of Green Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.2.11: Some Things to Know About Chemistry before You Even Start - *CC BY-NC-SA 4.0*
  - 2.5.2.12: Combining Atoms to Make Molecules and Compounds - *CC BY-NC-SA 4.0*
  - 2.5.2.13: The process of making and breaking chemical bonds- chemical reactions - *CC BY-NC-SA 4.0*
  - 2.5.2.14: The Nature of Matter and States of Matter - *CC BY-NC-SA 4.0*
  - 2.5.2.15: References and Questions - *CC BY-NC-SA 4.0*
- 2.5.3: The Elements - Basic Building Blocks of Green Chemicals - *CC BY-NC-SA 4.0*
  - 2.5.3.1: Elements, Atoms, and Atomic Theory - *CC BY-NC-SA 4.0*
  - 2.5.3.2: Hydrogen - The Simplest Atom - *CC BY-NC-SA 4.0*
  - 2.5.3.3: Helium - The First Noble Gas - *CC BY-NC-SA 4.0*
  - 2.5.3.4: Lithium, The First Metal - *CC BY-NC-SA 4.0*
  - 2.5.3.5: The Second Period of the Periodic Table - *CC BY-NC-SA 4.0*
  - 2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons - *CC BY-NC-SA 4.0*
  - 2.5.3.7: Completing the 20-Element Periodic Table - *CC BY-NC-SA 4.0*
  - 2.5.3.8: The Brief Periodic Table is Complete - *CC BY-NC-SA 4.0*
  - 2.5.3.E: The Elements - Basic Building Blocks of Green Chemicals (Exercises) - *CC BY-NC-SA 4.0*
  - 2.5.3.R: References - *CC BY-NC-SA 4.0*
- 2.5.4: Compounds- Safer Materials for a Safer World - *CC BY-NC-SA 4.0*
  - 2.5.4.1: Chemical Bonds and Compound Formation - *CC BY-NC-SA 4.0*
  - 2.5.4.2: Electrons Involved in Chemical Bonds and Octets of Electrons - *CC BY-NC-SA 4.0*
  - 2.5.4.3: Sodium Chloride and Ionic Bonds - *CC BY-NC-SA 4.0*
  - 2.5.4.4: Covalent Bonds in H2 and Other Molecules - *CC BY-NC-SA 4.0*
  - 2.5.4.5: Covalent Bonds in Compounds - *CC BY-NC-SA 4.0*
  - 2.5.4.6: Covalent Bonds and Green Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.4.7: Predicting Formulas of Covalently Bound Compounds - *CC BY-NC-SA 4.0*
  - 2.5.4.8: Chemical Formulas, the Mole, and Percentage Composition - *CC BY-NC-SA 4.0*
  - 2.5.4.9: What Are Chemical Compounds Called? - *CC BY-NC-SA 4.0*
  - 2.5.4.10: Acids, Bases and Salts - *CC BY-NC-SA 4.0*
  - 2.5.4.11: Questions and Problems - *CC BY-NC-SA 4.0*
- 2.5.5: Chemical Reactions- Making Materials Safely and Sustainable - *CC BY-NC-SA 4.0*
  - 2.5.5.1: Describing What Happens with Chemical Equations - *CC BY-NC-SA 4.0*
  - 2.5.5.2: Balancing Chemical Equations - *CC BY-NC-SA 4.0*
  - 2.5.5.3: Just Because You Can Write It Doesn't Mean That It Will Happen - *CC BY-NC-SA 4.0*
  - 2.5.5.4: Yield And Atom Economy in Chemical Reactions - *CC BY-NC-SA 4.0*
  - 2.5.5.5: Catalysts That Make Reactions Go - *CC BY-NC-SA 4.0*

- 2.5.5.6: Kinds of Chemical Reactions - *CC BY-NC-SA 4.0*
- 2.5.5.7: Oxidation-Reduction Reactions and Green Chemistry - *CC BY-NC-SA 4.0*
- 2.5.5.8: Quantitative Information from Chemical Reactions - *CC BY-NC-SA 4.0*
- 2.5.5.9: Energy in Chemical Reactions - *CC BY-NC-SA 4.0*
- 2.5.5.10: Stoichiometry by the Mole Ratio Method - *CC BY-NC-SA 4.0*
- 2.5.5.11: Limiting Reactant and Percent Yield - *CC BY-NC-SA 4.0*
- 2.5.5.12: Titrations - Measuring Moles by Volume of Solution - *CC BY-NC-SA 4.0*
- 2.5.5.13: Industrial Chemical Reactions - The Solvay Process - *CC BY-NC-SA 4.0*
- 2.5.5.14: Questions and Problems - *CC BY-NC-SA 4.0*

- 2.5.6: The Wonderful World of Carbon - Organic Chemistry and Biochemicals - *CC BY-NC-SA 4.0*
  - 2.5.6.1: Rings and Chains of Carbon Atoms - *CC BY-NC-SA 4.0*
  - 2.5.6.2: Compounds of Carbon and Hydrogen - Hydrocarbons - *CC BY-NC-SA 4.0*
  - 2.5.6.3: Using Lines To Show Organic Structural Formulas - *CC BY-NC-SA 4.0*
  - 2.5.6.4: Functional Groups - *CC BY-NC-SA 4.0*
  - 2.5.6.5: Giant Molecules from Small Organic Molecules - *CC BY-NC-SA 4.0*
  - 2.5.6.6: Questions and Problems - *CC BY-NC-SA 4.0*
  - 2.5.6.7: Supplementary References - *CC BY-NC-SA 4.0*
- 2.5.7: Chemistry of Life and Green Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.7.1: Green Biochemistry - *CC BY-NC-SA 4.0*
  - 2.5.7.2: Biochemistry and the Cells - *CC BY-NC-SA 4.0*
  - 2.5.7.3: Carbohydrates - *CC BY-NC-SA 4.0*
  - 2.5.7.4: Proteins - *CC BY-NC-SA 4.0*
  - 2.5.7.5: Lipids - Fats, Oils, and Hormones - *CC BY-NC-SA 4.0*
  - 2.5.7.6: Nucleic Acids - *CC BY-NC-SA 4.0*
  - 2.5.7.7: Enzymes - *CC BY-NC-SA 4.0*
  - 2.5.7.8: Biochemical Processes in Metabolism - *CC BY-NC-SA 4.0*
  - 2.5.7.9: Biochemistry of Toxic Substances and Toxicological Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.7.10: Literature Cited - *CC BY-NC-SA 4.0*
  - 2.5.7.11: Questions and Problems - *CC BY-NC-SA 4.0*
  - 2.5.7.12: Supplementary References - *CC BY-NC-SA 4.0*
- 2.5.8: The Five Environmental Spheres and Biogeochemical Cycle - *CC BY-NC-SA 4.0*
  - 2.5.8.1: Five Environmental Spheres - *CC BY-NC-SA 4.0*
  - 2.5.8.2: The Hydrosphere - *CC BY-NC-SA 4.0*
  - 2.5.8.3: The Atmosphere - *CC BY-NC-SA 4.0*
  - 2.5.8.4: The Geosphere - *CC BY-NC-SA 4.0*
  - 2.5.8.5: The Biosphere - *CC BY-NC-SA 4.0*
  - 2.5.8.6: The Anthrosphere - *CC BY-NC-SA 4.0*
  - 2.5.8.7: Cycles of Matter - *CC BY-NC-SA 4.0*
  - 2.5.8.8: Questions and Problems - *CC BY-NC-SA 4.0*
  - 2.5.8.9: Supplementary References - *CC BY-NC-SA 4.0*
- 2.5.9: Water - the Ultimate Green Substance - *CC BY-NC-SA 4.0*
  - 2.5.9.1: H_2O - Simple Formula, Remarkable Molecule - *CC BY-NC-SA 4.0*
  - 2.5.9.2: Occurrence, Availability , and Utilization of Water on Earth - *CC BY-NC-SA 4.0*
  - 2.5.9.3: Water Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.9.4: Water Pollution - *CC BY-NC-SA 4.0*
  - 2.5.9.5: Greening of Water - Purification Before Use - *CC BY-NC-SA 4.0*
  - 2.5.9.6: Wastewater Treatment - *CC BY-NC-SA 4.0*
  - 2.5.9.7: Advanced Wastewater Treatment and Recycling - *CC BY-NC-SA 4.0*
  - 2.5.9.8: The Many Uses of Water - *CC BY-NC-SA 4.0*
  - 2.5.9.9: Hot Water - Pressurized Subcritical Water - *CC BY-NC-SA 4.0*
  - 2.5.9.10: Supercritical Water - *CC BY-NC-SA 4.0*
  - 2.5.9.11: Literature Cited - *CC BY-NC-SA 4.0*
  - 2.5.9.12: Questions and Problems - *CC BY-NC-SA 4.0*
  - 2.5.9.13: Supplementary References - *CC BY-NC-SA 4.0*
- 2.5.10: Blue Skies for a Green Environment - *CC BY-NC-SA 4.0*
  - 2.5.10.1: More Than Just Air to Breathe - *CC BY-NC-SA 4.0*

- 2.5.10.2: The Gas Laws - *CC BY-NC-SA 4.0*
- 2.5.10.3: The Protective Atmosphere - *CC BY-NC-SA 4.0*
- 2.5.10.4: Atmospheric Chemistry and Photochemical Reactions - *CC BY-NC-SA 4.0*
- 2.5.10.5: The Enormous Importance of Climate - *CC BY-NC-SA 4.0*
- 2.5.10.6: Atmospheric Particle Pollutants - *CC BY-NC-SA 4.0*
- 2.5.10.7: Pollutant Gaseous Oxides - *CC BY-NC-SA 4.0*
- 2.5.10.8: Acid Rain - *CC BY-NC-SA 4.0*
- 2.5.10.9: Miscellaneous Gases in the Atmosphere - *CC BY-NC-SA 4.0*
- 2.5.10.10: Photochemical Smog - *CC BY-NC-SA 4.0*
- 2.5.10.11: Natural Capital of the Atmosphere - *CC BY-NC-SA 4.0*
- 2.5.10.12: Energy and Mass Transfer in the Atmosphere - *CC BY-NC-SA 4.0*
- 2.5.10.13: Questions and Problems - *Undeclared*
- 2.5.10.14: Supplementary References - *Undeclared*

- 2.5.11: The Geosphere and a Green Earth - *CC BY-NC-SA 4.0*
  - 2.5.11.1: Probing Deep into the Geosphere - A Well Too Far? - *CC BY-NC-SA 4.0*
  - 2.5.11.2: The Nature of the Geosphere - *CC BY-NC-SA 4.0*
  - 2.5.11.3: The Geosphere as a Source of Natural Capital - *CC BY-NC-SA 4.0*
  - 2.5.11.4: Environmental Hazards of the Geosphere - *CC BY-NC-SA 4.0*
  - 2.5.11.5: Water on and in the Geosphere - *CC BY-NC-SA 4.0*
  - 2.5.11.6: Anthrospheric Influences on the Geosphere - *CC BY-NC-SA 4.0*
  - 2.5.11.7: The Geosphere as a Waste Repository - *CC BY-NC-SA 4.0*
  - 2.5.11.8: Have You Thanked a Clod Today? - *CC BY-NC-SA 4.0*
  - 2.5.11.9: Plant Nutrients and Fertilizers in Soil - *CC BY-NC-SA 4.0*
  - 2.5.11.10: Soil and Plants Related To Wastes and Pollutants - *CC BY-NC-SA 4.0*
  - 2.5.11.11: Soil Loss - Desertification and Deforestation - *CC BY-NC-SA 4.0*
  - 2.5.11.12: Literature Cited and Supplementary References - *Undeclared*
  - 2.5.11.13: Questions and Problems - *Undeclared*
- 2.5.12: The Biosphere and the Role of Green Chemistry in Feeding a Hungry World - *CC BY-NC-SA 4.0*
  - 2.5.12.1: Pigweed's Revenge - *CC BY-NC-SA 4.0*
  - 2.5.12.2: The Biosphere - *CC BY-NC-SA 4.0*
  - 2.5.12.3: Cells - Basic Units of Life - *CC BY-NC-SA 4.0*
  - 2.5.12.4: Metabolism and Control in Organisms - *CC BY-NC-SA 4.0*
  - 2.5.12.5: Reproduction and Inherited Traits - *CC BY-NC-SA 4.0*
  - 2.5.12.6: Stability and Equilibrium of the Biosphere - *CC BY-NC-SA 4.0*
  - 2.5.12.7: DNA and the Human Genome - *CC BY-NC-SA 4.0*
  - 2.5.12.8: Genetic Engineering - *CC BY-NC-SA 4.0*
  - 2.5.12.9: Biological Interaction with Environmental Chemicals - *CC BY-NC-SA 4.0*
  - 2.5.12.10: Biodegradation - *CC BY-NC-SA 4.0*
  - 2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture - *CC BY-NC-SA 4.0*
  - 2.5.12.12: Agricultural Applications of Genetically Modified Organisms - *CC BY-NC-SA 4.0*
  - 2.5.12.13: The Anthrosphere in Support of the Biosphere - *CC BY-NC-SA 4.0*
  - 2.5.12.14: Livestock and their Wastes - *CC BY-NC-SA 4.0*
  - 2.5.12.15: Questions and Problems - *Undeclared*
  - 2.5.12.16: Supplementary References - *Undeclared*
- 2.5.13: The Anthrosphere, Industrial Ecology, and Green Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.13.1: Industrial Ecology and Industrial Ecosystems - *CC BY-NC-SA 4.0*
  - 2.5.13.2: Metabolic Processes in Industrial Ecosystems - *CC BY-NC-SA 4.0*
  - 2.5.13.3: Life Cycles in Industrial Ecosystems - *CC BY-NC-SA 4.0*
  - 2.5.13.4: Kinds of Products - *CC BY-NC-SA 4.0*

- 2.5.13.5: Attributes Required by an Industrial Ecosystem - *CC BY-NC-SA 4.0*
- 2.5.13.6: The Kalundborg Industrial Ecosystem - *CC BY-NC-SA 4.0*
- 2.5.13.7: Environmental Impacts of Industrial Ecosystems - *CC BY-NC-SA 4.0*
- 2.5.13.8: Green Chemistry and Industrial Ecology - *CC BY-NC-SA 4.0*
- 2.5.13.9: Predicting and Reducing Hazards - *CC BY-NC-SA 4.0*
- 2.5.13.10: The E-Factor in Green Chemistry - *CC BY-NC-SA 4.0*
- 2.5.13.11: Catalysts and Catalysis - *CC BY-NC-SA 4.0*
- 2.5.13.12: Biocatalysis with Enzymes - *CC BY-NC-SA 4.0*
- 2.5.13.13: Energizing Chemical Reactions and Process Intensification - *CC BY-NC-SA 4.0*
- 2.5.13.14: Solvents and Alternate Reaction Media - *CC BY-NC-SA 4.0*
- 2.5.13.15: Feedstocks and Reagents - *CC BY-NC-SA 4.0*
- 2.5.13.16: Literature Cited and Supplementary References - *Undeclared*
- 2.5.13.17: Questions and Problems - *Undeclared*

- 2.5.14: Feeding the Anthrosphere- Utilizing Renewable and Biological Materials - *CC BY-NC-SA 4.0*
  - 2.5.14.1: Feeding the Anthrosphere - *CC BY-NC-SA 4.0*
  - 2.5.14.2: Utilization of Feedstocks - *CC BY-NC-SA 4.0*
  - 2.5.14.3: Biological Feedstocks - *CC BY-NC-SA 4.0*
  - 2.5.14.4: Biological Sources of Chemicals - *CC BY-NC-SA 4.0*
  - 2.5.14.5: Biorefineries and Biomass Utilization - *CC BY-NC-SA 4.0*
  - 2.5.14.6: Monosaccharide Feedstocks - Glucose, Fructose, and Xylose - *CC BY-NC-SA 4.0*
  - 2.5.14.7: Cellulose - *CC BY-NC-SA 4.0*
  - 2.5.14.8: Feedstocks from Cellulose Wastes - *CC BY-NC-SA 4.0*
  - 2.5.14.9: Lignin - *CC BY-NC-SA 4.0*
  - 2.5.14.10: Direct Biosynthesis of Polymers - *CC BY-NC-SA 4.0*
  - 2.5.14.11: Bioconversion Processes for Synthetic Chemicals - *Undeclared*
  - 2.5.14.12: Bamboo - Ancient Material for the Future - *CC BY-NC-SA 4.0*
  - 2.5.14.13: Literature Cited - *Undeclared*
  - 2.5.14.14: Questions and Problems - *Undeclared*
- 2.5.15: Sustainable Energy- The Essential Basis of Green Systems - *CC BY-NC-SA 4.0*
  - 2.5.15.1: Sustainability and Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.2: What is Energy? - *CC BY-NC-SA 4.0*
  - 2.5.15.3: Radiant Energy from the Sun - *CC BY-NC-SA 4.0*
  - 2.5.15.4: Sources of Energy Used in the Anthrosphere - *CC BY-NC-SA 4.0*
  - 2.5.15.5: Conversions Between Forms of Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.6: Green Technology for Energy Conversion - *CC BY-NC-SA 4.0*
  - 2.5.15.7: Energy Conservation - *CC BY-NC-SA 4.0*
  - 2.5.15.8: Depletable Fossil Fuels - *CC BY-NC-SA 4.0*
  - 2.5.15.9: Carbon Sequestration for Fossil Fuel Utilization - *CC BY-NC-SA 4.0*
  - 2.5.15.10: Nuclear Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.11: Renewable Energy Sources - Solar Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.12: Energy from Wind and Water - *CC BY-NC-SA 4.0*
  - 2.5.15.13: Biomass Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.14: Geothermal Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.15: Hydrogen for Energy Storage and Utilization - *CC BY-NC-SA 4.0*
  - 2.5.15.16: Literature Cited and Supplementary References - *Undeclared*
  - 2.5.15.17: Questions and Problems - *Undeclared*
- 2.5.16: Terrorism, Toxicity, and Vulnerability- Green Chemistry and Technology in Defense of Human Welfare - *CC BY-NC-SA 4.0*
  - 2.5.16.1: Vulnerability to Terrorist Attack - *CC BY-NC-SA 4.0*
  - 2.5.16.2: Protecting the Anthrosphere - *CC BY-NC-SA 4.0*

- 2.5.16.3: Substances That Explode, Burn, or React Violently - *CC BY-NC-SA 4.0*
- 2.5.16.4: Toxic Substances and Toxicology - *CC BY-NC-SA 4.0*
- 2.5.16.5: Toxic Chemical Attack - *CC BY-NC-SA 4.0*
- 2.5.16.6: Protecting Water, Food, and Air - *CC BY-NC-SA 4.0*
- 2.5.16.7: Detecting Hazards - *CC BY-NC-SA 4.0*
- 2.5.16.8: Green Chemistry to Combat Terrorism - *CC BY-NC-SA 4.0*
- 2.5.16.9: Green Chemistry for Sustainable Prosperity and a Safer World - *CC BY-NC-SA 4.0*
- 2.5.16.10: Green Chemistry to Combat Terrorism - *Undeclared*
- 2.5.16.11: Green Chemistry for Sustainable Prosperity and and a Safer World - *Undeclared*
- 2.5.16.12: Literature Cited - *Undeclared*
- 2.5.16.13: Questions and Problems - *Undeclared*

- 2.5.17: The Ten Commandments of Sustainability and Sensible Measure - *CC BY-NC-SA 4.0*
  - 2.5.17.1: We Cannot Go on Like This - *CC BY-NC-SA 4.0*
  - 2.5.17.2: The First Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.3: The Second Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.4: The Third Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.5: The Fourth Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.6: The Fifth Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.7: The Sixth Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.8: The Seventh Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.9: The Ninth Commandment - *Undeclared*
  - 2.5.17.10: The Tenth Commandment - *Undeclared*
  - 2.5.17.11: Some Sensible Measures for Sustainability - *Undeclared*
  - 2.5.17.12: The Eighth Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.13: Literature Cited - *Undeclared*
  - 2.5.17.14: Questions and Problems - *Undeclared*
  - 2.5.17.15: Supplementary References - *Undeclared*
- Back Matter - *CC BY-NC-SA 4.0*
  - Detailed Licensing - *Undeclared*

- 2.6: Toxicology MSDT - *CC BY-NC 4.0*
  - Front Matter - *CC BY-NC 4.0*
    - Licensing - *Undeclared*
    - Preface - *CC BY-NC 4.0*
  - 2.6.1: Pathophysiology - *CC BY-NC 4.0*
    - 2.6.1.1: What is Pathophysiology? - *CC BY-NC 4.0*
    - 2.6.1.2: Outcomes of Targeted and Non-Targeted Toxicity - *CC BY-NC 4.0*
    - 2.6.1.3: Cellular Response to Toxicant-Induced Injury - *CC BY-NC 4.0*
    - 2.6.1.4: Repair and Adaptation - *CC BY-NC 4.0*
    - 2.6.1.5: Patterns of Toxic Injury - *CC BY-NC 4.0*
    - 2.6.1.6: Final Evaluation - *CC BY-NC 4.0*
  - 2.6.2: Biochemistry and Molecular Genetics - *CC BY-NC 4.0*
    - 2.6.2.1: Introduction to Biomolecules and Cell Components - *CC BY-NC 4.0*
    - 2.6.2.2: Cell Structure and Subcellular Compartments - *CC BY-NC 4.0*
    - 2.6.2.3: DNA and RNA Metabolism - *CC BY-NC 4.0*
    - 2.6.2.4: Epigenetic Mechanisms - *CC BY-NC 4.0*
    - 2.6.2.5: Section 2 Final Evaluation - *CC BY-NC 4.0*
  - 2.6.3: Principles of Genetic Toxicology - *CC BY-NC 4.0*
    - 2.6.3.1: Introduction to Genetic-toxicology Assay - *CC BY-NC 4.0*
    - 2.6.3.2: Different Genetic Damages or Mutations - *CC BY-NC 4.0*
    - 2.6.3.3: Different Genetic-Toxicology Assays - *CC BY-NC 4.0*
    - 2.6.3.4: Different Cytotoxicity Assays - *CC BY-NC 4.0*
    - 2.6.3.5: Epigenetics Assay - *CC BY-NC 4.0*
    - 2.6.3.6: Section 3 Final Evaluation - *CC BY-NC 4.0*

- 2.6.4: Applied Systems Toxicology - *CC BY-NC 4.0*
  - 2.6.4.1: Systems Toxicology - *CC BY-NC 4.0*
  - 2.6.4.2: Dose Level and Applied Toxicology - *CC BY-NC 4.0*
  - 2.6.4.3: Tools and Technologies in Systems Toxicology - *CC BY-NC 4.0*
  - 2.6.4.4: Other Approaches for Predictive Toxicity Modeling - *CC BY-NC 4.0*
  - 2.6.4.5: Technologies Used In Systems Biology/Toxicology - *CC BY-NC 4.0*
  - 2.6.4.6: Takeaways Summary - *CC BY-NC 4.0*
  - 2.6.4.7: Section 4 Final Evaluation - *CC BY-NC 4.0*
- 2.6.5: Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.1: Introduction to Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.2: Global Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.3: Topic 3- Regional Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.4: National Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.5: State Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.6: Non-Governmental Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.7: Section 5 Final Evaluation - *CC BY-NC 4.0*
- 2.6.6: Principles of Toxicology - *CC BY-NC 4.0*
  - 2.6.6.1: Introduction to Toxicology - *CC BY-NC 4.0*
    - 2.6.6.1.1: What is toxicology? - *CC BY-NC 4.0*
    - 2.6.6.1.2: Basic Terminology - *CC BY-NC 4.0*
  - 2.6.6.2: Dose and Dose Response - *CC BY-NC 4.0*
    - 2.6.6.2.1: Dose and It's Impact on Toxicity - *CC BY-NC 4.0*
    - 2.6.6.2.2: The Dose Response Relationship - *CC BY-NC 4.0*
    - 2.6.6.2.3: Dose Estimates of Toxic Effects - *CC BY-NC 4.0*
    - 2.6.6.2.4: Determining the Safety of a Drug - *CC BY-NC 4.0*
    - 2.6.6.2.5: NOAEL and LOAEL - *CC BY-NC 4.0*
  - 2.6.6.3: Toxic Effects - *CC BY-NC 4.0*
    - 2.6.6.3.1: Types of Toxic Effects - *CC BY-NC 4.0*
    - 2.6.6.3.2: Factors Affecting Toxicity - *CC BY-NC 4.0*
    - 2.6.6.3.3: Systemic Toxic Effects - *CC BY-NC 4.0*
    - 2.6.6.3.4: Organ Specific Toxic Effects - *CC BY-NC 4.0*
  - 2.6.6.4: Interactions - *CC BY-NC 4.0*
    - 2.6.6.4.1: Interactions - *CC BY-NC 4.0*
  - 2.6.6.5: Toxicity Testing Methods - *CC BY-NC 4.0*
    - 2.6.6.5.1: Testing and Assessing Toxicity - *CC BY-NC 4.0*
    - 2.6.6.5.2: Clinical Investigations and Other Types of Human Data - *CC BY-NC 4.0*
    - 2.6.6.5.3: Epidemiology Studies - *CC BY-NC 4.0*
  - 2.6.6.6: Risk Assessment - *CC BY-NC 4.0*
    - 2.6.6.6.1: Risk Assessment - *CC BY-NC 4.0*
    - 2.6.6.6.2: Hazard Identification - *CC BY-NC 4.0*
    - 2.6.6.6.3: Dose-Response Assessment - *CC BY-NC 4.0*
    - 2.6.6.6.4: Exposure Assessment - *CC BY-NC 4.0*
    - 2.6.6.6.5: Risk Characterization - *CC BY-NC 4.0*
  - 2.6.6.7: Exposure Standards and Guidelines - *CC BY-NC 4.0*
    - 2.6.6.7.1: Exposure Standards and Guidelines - *CC BY-NC 4.0*
    - 2.6.6.7.2: Regulation of Consumer Products and Drug Safety - *CC BY-NC 4.0*
    - 2.6.6.7.3: Environmental Exposure Standards/Guidelines - *CC BY-NC 4.0*
    - 2.6.6.7.4: Occupational (Workplace) Exposure Standards/Guidelines/Approaches - *CC BY-NC 4.0*
  - 2.6.6.8: Basic Physiology - *CC BY-NC 4.0*
    - 2.6.6.8.1: Introduction to Basic Physiology - *CC BY-NC 4.0*
    - 2.6.6.8.2: Homeostasis - *CC BY-NC 4.0*
    - 2.6.6.8.3: Organs and Organ Systems - *CC BY-NC 4.0*
    - 2.6.6.8.4: Tissues - *CC BY-NC 4.0*

- 2.6.6.8.5: Cells - *CC BY-NC 4.0*
- 2.6.6.8.6: Chemicals - *CC BY-NC 4.0*
- 2.6.6.9: Introduction to Toxicokinetics - *CC BY-NC 4.0*
  - 2.6.6.9.1: What is Toxicokinetics - *CC BY-NC 4.0*
- 2.6.6.10: Absorption - *CC BY-NC 4.0*
  - 2.6.6.10.1: Introduction to Absorption - *CC BY-NC 4.0*
  - 2.6.6.10.2: Gastrointestinal Tract - *CC BY-NC 4.0*
  - 2.6.6.10.3: Respiratory Tract - *CC BY-NC 4.0*
  - 2.6.6.10.4: Dermal Route - *CC BY-NC 4.0*
  - 2.6.6.10.5: Other Routes of Exposure - *CC BY-NC 4.0*
- 2.6.6.11: Distribution - *CC BY-NC 4.0*
  - 2.6.6.11.1: Introduction to Distribution - *CC BY-NC 4.0*
  - 2.6.6.11.2: Influence of Route of Exposure - *CC BY-NC 4.0*
  - 2.6.6.11.3: Disposition Models - *CC BY-NC 4.0*
  - 2.6.6.11.4: Structural Barriers to Distribution - *CC BY-NC 4.0*
  - 2.6.6.11.5: Storage Sites - *CC BY-NC 4.0*
- 2.6.6.12: Biotransformation - *CC BY-NC 4.0*
  - 2.6.6.12.1: Introduction to Biotransformation - *CC BY-NC 4.0*
  - 2.6.6.12.2: Chemical Reactions - *CC BY-NC 4.0*
  - 2.6.6.12.3: Biotransformation Sites - *CC BY-NC 4.0*
  - 2.6.6.12.4: Modifiers of Biotransformation - *CC BY-NC 4.0*
- 2.6.6.13: Excretion - *CC BY-NC 4.0*
  - 2.6.6.13.1: Introduction to Secretion - *CC BY-NC 4.0*
  - 2.6.6.13.2: Urinary Excretion - *CC BY-NC 4.0*
  - 2.6.6.13.3: Fecal Excretion - *CC BY-NC 4.0*
  - 2.6.6.13.4: Exhaled Air - *CC BY-NC 4.0*
  - 2.6.6.13.5: Other Routes - *CC BY-NC 4.0*
- 2.6.6.14: Cellular Toxicology - *CC BY-NC 4.0*
  - 2.6.6.14.1: Adaptation - *CC BY-NC 4.0*
  - 2.6.6.14.2: Cell Damage and Tissue Repair - *CC BY-NC 4.0*
  - 2.6.6.14.3: Cancer - *CC BY-NC 4.0*
  - 2.6.6.14.4: Neurotoxicity - *CC BY-NC 4.0*
- 2.6.6.15: Intuitive Toxicology and Risk Communication - *CC BY-NC 4.0*
  - 2.6.6.15.1: Intuitive Toxicology - *CC BY-NC 4.0*
  - 2.6.6.15.2: Risk Communication - *CC BY-NC 4.0*
- 2.6.6.16: Environmental Toxicology, Environmental Health, and One Health - *CC BY-NC 4.0*
  - 2.6.6.16.1: Environmental Toxicology - *CC BY-NC 4.0*
  - 2.6.6.16.2: Environmental Health - *CC BY-NC 4.0*
  - 2.6.6.16.3: One Health - *CC BY-NC 4.0*
- 2.6.6.17: Conclusion - *CC BY-NC 4.0*
- Back Matter - *CC BY-NC 4.0*
  - Detailed Licensing - *Undeclared*
- Back Matter - *Undeclared*
  - Index - *Undeclared*
  - Glossary - *Undeclared*
  - Detailed Licensing - *Undeclared*

3: Ecology - *CC BY-NC 4.0*

- 3.1: Ecology Overview - *CC BY-NC 4.0*
- 3.2: Populations - *CC BY-NC 4.0*
  - 3.2.1: Population Dispersion - *CC BY-NC 4.0*
  - 3.2.2: Population Size - *CC BY-NC 4.0*
  - 3.2.3: Population Growth and Regulation - *CC BY-SA 4.0*
  - 3.2.4: Life History - *CC BY-NC 4.0*
  - 3.2.5: Data Dive- Wildlife Corridors - *Undeclared*
  - 3.2.6: Review - *CC BY-NC 4.0*
- 3.3: Communities - *CC BY-NC 4.0*
  - 3.3.1: Biotic Interactions - *CC BY-NC 4.0*
    - 3.3.1.1: Trophic Interactions - *CC BY-NC 4.0*
      - 3.3.1.1.1: Predation - *CC BY-SA 4.0*
      - 3.3.1.1.2: Herbivory - *CC BY-SA 4.0*
      - 3.3.1.1.3: Parasitism - *CC BY-NC 4.0*
      - 3.3.1.1.4: Food Chains and Food Webs - *CC BY-NC 4.0*
    - 3.3.1.2: Competition - *CC BY-SA 4.0*
    - 3.3.1.3: Facilitation - *CC BY-NC 4.0*
    - 3.3.1.4: Symbiosis - *CC BY-NC 4.0*
  - 3.3.2: Community Structure and Dynamics - *CC BY-NC 4.0*

- 3.3.3: Data Dive- Bison Impacts In Prairies - *Undeclared*
- 3.3.4: Review - *CC BY-NC 4.0*
- 3.4: Ecosystems - *CC BY-NC 4.0*
  - 3.4.1: Ecosystem Types and Dynamics - *CC BY-NC 4.0*
  - 3.4.2: Matter - *CC BY-NC 4.0*
  - 3.4.3: Biogeochemical Cycles - *CC BY-SA 4.0*
  - 3.4.4: Soils - *CC BY-SA 4.0*
  - 3.4.5: Soil Degradation - *CC BY-NC-SA 4.0*
  - 3.4.6: Data Dive- Beaver Impacts on Wetlands - *Undeclared*
  - 3.4.7: Review - *CC BY-NC 4.0*
- 3.5: Biomes - *CC BY-SA 4.0*
  - 3.5.1: Climate and Biomes - *CC BY-NC-SA 4.0*
  - 3.5.2: Terrestrial Biomes - *CC BY-NC 4.0*
  - 3.5.3: Aquatic Biomes - *CC BY-SA 4.0*
  - 3.5.4: Data Dive- Biome Carbon Storage - *Undeclared*
  - 3.5.5: Review - *CC BY 4.0*

4: Conservation - *CC BY-NC 4.0*

- 4.1: The Value of Biodiversity - *CC BY-NC 4.0*
  - 4.1.1: Ecosystem Diversity - *CC BY-SA 4.0*
  - 4.1.2: Species Diversity - *CC BY-NC 4.0*
  - 4.1.3: Genetic Diversity - *CC BY-SA 4.0*
  - 4.1.4: Patterns of Biodiversity - *CC BY-SA 4.0*
  - 4.1.5: Data Dive- Biodiversity and Drugs - *Undeclared*
  - 4.1.6: Review - *CC BY-NC 4.0*
- 4.2: Threats to Biodiversity - *CC BY-NC 4.0*
  - 4.2.1: Extinction - *CC BY-NC 4.0*
  - 4.2.2: Measures of Biodiversity Loss - *CC BY-SA 4.0*
  - 4.2.3: Habitat Loss - *CC BY-SA 4.0*
  - 4.2.4: Overexploitation - *CC BY-SA 4.0*
  - 4.2.5: Pollution - *CC BY-SA 4.0*
  - 4.2.6: Invasive Species - *CC BY-NC 4.0*
  - 4.2.7: Climate Change - *CC BY-NC 4.0*
  - 4.2.8: Data Dive- Threats to Biodiversity - *Undeclared*
  - 4.2.9: Review - *CC BY-NC 4.0*
- 4.3: Protecting Biodiveristy - *CC BY-NC 4.0*
  - 4.3.1: Policies - *CC BY-NC 4.0*
  - 4.3.2: Non-profit Organizations - *CC BY-SA 4.0*
  - 4.3.3: Species-level Conservation - *CC BY-NC 4.0*
  - 4.3.4: Protected Areas - *CC BY-NC 4.0*
  - 4.3.5: Ecosystem Restoration - *CC BY-NC 4.0*
  - 4.3.6: Economic Influences on Conservation - *CC BY-SA 4.0*
  - 4.3.7: Individual Choices - *CC BY-NC 4.0*
  - 4.3.8: Data Dive- Island Fox Populations - *Undeclared*
  - 4.3.9: Review - *CC BY-NC 4.0*

5: Humans and the Environment - *CC BY-NC 4.0*

- 5.1: The Human Population - *CC BY-NC-SA 4.0*
  - 5.1.1: History of Human Population Growth - *CC BY-NC-SA 4.0*
  - 5.1.2: The Rate of Human Population Growth - *CC BY-SA 4.0*
  - 5.1.3: Looking Ahead - *CC BY-NC 4.0*
  - 5.1.4: Data Dive- World Population Densities - *Undeclared*
  - 5.1.5: Review - *CC BY-NC 4.0*
- 5.2: Water Resources - *CC BY-NC 4.0*
  - 5.2.1: Fresh Water Supply and the Water Cycle - *CC BY-NC-SA 4.0*
  - 5.2.2: Water Usage - *CC BY-NC-SA 4.0*
  - 5.2.3: Water Scarcity and Solutions - *CC BY-SA 4.0*
  - 5.2.4: Data Dive- Aqueducts of Rome - *Undeclared*
  - 5.2.5: Review - *CC BY-NC 4.0*
- 5.3: Agriculture - *CC BY-NC 4.0*
  - 5.3.1: Food Security - *CC BY-SA 4.0*
  - 5.3.2: Industrial Agriculture - *CC BY-SA 4.0*
  - 5.3.3: Selective Breeding and Genetic Engineering - *CC BY-SA 4.0*
  - 5.3.4: Sustainable Agriculture - *CC BY-SA 4.0*
  - 5.3.5: Data Dive- Landslides and Ag Terraces - *Undeclared*
  - 5.3.6: Review - *CC BY-NC 4.0*
- 5.4: Environmental Health - *CC BY 4.0*
  - 5.4.1: Types of Environmental Hazards - *CC BY-SA 4.0*
  - 5.4.2: Epidemiology - *CC BY-SA 4.0*
  - 5.4.3: Infectious Diseases - *CC BY-NC 4.0*
  - 5.4.4: Environmental Toxicology - *CC BY-SA 4.0*
  - 5.4.5: Environmental Hazard Reduction - *CC BY-SA 4.0*
  - 5.4.6: Data Dive- Cholera Cases Worldwide - *Undeclared*
  - 5.4.7: Review - *CC BY-NC 4.0*

6: Energy - *CC BY-NC 4.0*

- 6.1: Fossil Fuels - *CC BY 4.0*

- 6.1.1: Types of Fossil Fuels and Formation - *CC BY-NC-SA 4.0*
- 6.1.2: Mining, Processing, and Generating Electricity - *CC BY-NC-SA 4.0*
- 6.1.3: Fossil Fuel Consumption - *CC BY-NC 4.0*
- 6.1.4: Consequences of Fossil Fuels - *CC BY-NC-SA 4.0*
- 6.1.5: Data Dive- Global Fossil Fuel Consumption - *Undeclared*
- 6.1.6: Review - *CC BY-NC 4.0*

- 6.2: Nuclear Energy - *CC BY-SA 4.0*
  - 6.2.1: Radioactive Isotopes - *CC BY-NC 4.0*
  - 6.2.2: Generating Electricity with Nuclear Energy - *CC BY-NC 4.0*
  - 6.2.3: Nuclear Energy Consumption - *CC BY-NC 4.0*
  - 6.2.4: Consequences of Nuclear Energy - *CC BY-SA 4.0*
  - 6.2.5: Data Dive- Global Nuclear Power Generation - *Undeclared*
  - 6.2.6: Review - *CC BY-NC 4.0*
- 6.3: Renewable Energy - *CC BY-NC 4.0*
  - 6.3.1: Renewable Energy History and Consumption - *CC BY-SA 4.0*
  - 6.3.2: Wind Energy - *CC BY-NC 4.0*
  - 6.3.3: Solar Energy - *CC BY-NC 4.0*
  - 6.3.4: Geothermal Energy - *CC BY-NC 4.0*
  - 6.3.5: Hydropower - *CC BY-NC 4.0*
  - 6.3.6: Biofuels (Biomass Energy) - *CC BY-NC 4.0*
  - 6.3.7: Energy Conservation - *CC BY-NC 4.0*
  - 6.3.8: Data Dive- Global Renewable Energy Generation - *Undeclared*
  - 6.3.9: Review - *CC BY-NC 4.0*

- 7: Environmental Impacts - *CC BY 4.0*
  - 7.1: Solid Waste Management - *CC BY-SA 4.0*
    - 7.1.1: Waste Generation - *CC BY-NC 4.0*
    - 7.1.2: Waste Disposal - *CC BY-NC 4.0*
    - 7.1.3: Solid Waste and Marine Life - *CC BY-NC 4.0*
    - 7.1.4: Waste Reduction - *CC BY-NC 4.0*
    - 7.1.5: Data Dive- "Flushable" Wipes - *Undeclared*
    - 7.1.6: Review - *CC BY-NC 4.0*
  - 7.2: Pollution - *Undeclared*
    - 7.2.1: Water Pollution - *CC BY-NC 4.0*
      - 7.2.1.1: Water Pollutants and Their Sources - *CC BY-NC-SA 4.0*
      - 7.2.1.2: Water Treatment - *CC BY-SA 4.0*
      - 7.2.1.3: Mitigating Water Pollution - *CC BY-NC-SA 4.0*
    - 7.2.2: Air Pollution - *CC BY-SA 4.0*
      - 7.2.2.1: Ozone Depletion - *CC BY-SA 4.0*
      - 7.2.2.2: Acid Deposition - *CC BY 4.0*
    - 7.2.3: Data Dive- Plastic in World Oceans - *Undeclared*
    - 7.2.4: Review - *CC BY 4.0*
  - 7.3: Climate Change - *CC BY 4.0*
    - 7.3.1: The Greenhouse Effect and Climate Change - *CC BY-SA 4.0*
    - 7.3.2: Data Dive- Flooding in Venice Italy - *Undeclared*
    - 7.3.3: Review - *CC BY 4.0*
- 8: Sustainability - *CC BY-NC 4.0*
  - 8.1: Introduction to Sustainability - *Undeclared*
  - 8.2: History of Sustainability - *CC BY-NC 4.0*
  - 8.3: Environmental Ethics - *CC BY-NC 4.0*
    - 8.3.1: Environmental Justice and Indigenous Struggles - *CC BY-NC 4.0*
    - 8.3.2: Perspectives of the Environment - *CC BY-NC 4.0*
  - 8.4: Environmental Policy - *Undeclared*
  - 8.5: Sustainable Cities - *CC BY-NC 4.0*
  - 8.6: Sustainability and the Future - *CC BY-NC 4.0*
  - 8.7: Data Dive- Sustainable Investments - *Undeclared*
  - 8.8: Review - *CC BY 4.0*
- Back Matter - *Undeclared*
  - Index - *Undeclared*
  - Glossary - *Undeclared*
  - Detailed Licensing - *Undeclared*

CHAPTER OVERVIEW

3: Ecology

Ecology is the study of the interactions of living organisms with their environment. One core goal of ecology is to understand the distribution and abundance of living things in the physical environment. Attainment of this goal requires the integration of scientific disciplines inside and outside of biology, such as biochemistry, physiology, evolution, biodiversity, molecular biology, geology, and climatology. Some ecological research also applies aspects of chemistry and physics, and it frequently uses mathematical models.

Why study ecology? Perhaps you are interested in learning about the natural world and how living things have adapted to the physical conditions of their environment. Or, perhaps you're a future physician seeking to understand the connection between human health and ecology.

Humans are a part of the ecological landscape, and human health is one important part of human interaction with our physical and living environment. Lyme disease, for instance, serves as one modern-day example of the connection between our health and the natural world. More formally known as Lyme borreliosis, Lyme disease is a bacterial infection that can be transmitted to humans when they are bitten by the deer tick (*Ixodes scapularis*), which is the primary vector for this disease (figure 3.a). However, not all deer ticks carry the bacteria that will cause Lyme disease in humans, and *I. scapularis* can have other hosts besides deer. In fact, it turns out that the probability of infection depends on the type of host upon which the tick develops: a higher proportion of ticks that live on white-footed mice carry the bacterium than do ticks that live on deer. Knowledge about the environments and population densities in which the host species is abundant would help a physician or an epidemiologist better understand how Lyme disease is transmitted and how its incidence could be reduced.



(a) (b) (c)

Figure 3.a: The (a) deer tick carries the bacterium that produces Lyme disease in humans, often evident in (b) a symptomatic bull's eye rash. The (c) white-footed mouse is one well-known host to deer ticks carrying the Lyme disease bacterium. (credit a: modification of work by Scott Bauer, USDA ARS; credit b: modification of work by James Gathany, CDC; credit c: modification of work by Rob Ireton)

Ecologists ask questions across four levels of biological organization—organismal, population, community, and ecosystem. At the organismal level, ecologists study individual organisms and how they interact with their environments. At the population and community levels, ecologists explore, respectively, how a population of organisms changes over time and the ways in which that population interacts with other species in the community. Ecologists studying an ecosystem examine the living species (the biotic components) of the ecosystem as well as the nonliving portions (the abiotic components), such as air, water, and soil, of the environment. The objective of this unit will be to further explore these topics and how they connect to and overlap with environmental science.

Attribution

Modified by Melissa Ha and Rachel Schleiger from Ecology and the Biosphere by OpenStax (licensed under CC-BY)

- 3.1: Ecology Overview
- 3.2: Populations
  - 3.2.1: Population Dispersion
  - 3.2.2: Population Size
  - 3.2.3: Population Growth and Regulation
  - 3.2.4: Life History
  - 3.2.5: Data Dive- Wildlife Corridors
  - 3.2.6: Review
- 3.3: Communities
  - 3.3.1: Biotic Interactions
    - 3.3.1.1: Trophic Interactions
      - 3.3.1.1.1: Predation
      - 3.3.1.1.2: Herbivory
      - 3.3.1.1.3: Parasitism
      - 3.3.1.1.4: Food Chains and Food Webs
    - 3.3.1.2: Competition
    - 3.3.1.3: Facilitation
    - 3.3.1.4: Symbiosis
  - 3.3.2: Community Structure and Dynamics
  - 3.3.3: Data Dive- Bison Impacts In Prairies
  - 3.3.4: Review
- 3.4: Ecosystems
  - 3.4.1: Ecosystem Types and Dynamics
  - 3.4.2: Matter
  - 3.4.3: Biogeochemical Cycles
  - 3.4.4: Soils
  - 3.4.5: Soil Degradation
  - 3.4.6: Data Dive- Beaver Impacts on Wetlands
  - 3.4.7: Review
- 3.5: Biomes
  - 3.5.1: Climate and Biomes
  - 3.5.2: Terrestrial Biomes
  - 3.5.3: Aquatic Biomes
  - 3.5.4: Data Dive- Biome Carbon Storage
  - 3.5.5: Review

This page titled 3: Ecology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

3.1: Ecology Overview

Levels of Ecological Study

When a discipline such as biology is studied, it is often helpful to subdivide it into smaller, related areas. For instance, cell biologists interested in cell signaling need to understand the chemistry of the signal molecules (which are usually proteins) as well as the result of cell signaling. Ecologists interested in the factors that influence the survival of an endangered species might use mathematical models to predict how current conservation efforts affect endangered organisms. To produce a sound set of management options, a conservation biologist needs to collect accurate data, including current population size, factors affecting reproduction (like physiology and behavior), habitat requirements (such as plants and soils), and potential human influences on the endangered population and its habitat (which might be derived through studies in sociology and urban ecology). Within the discipline of ecology, researchers work at four specific levels, sometimes discretely and sometimes with overlap: organism, population, community, and ecosystem (figure 3.1.*a*).

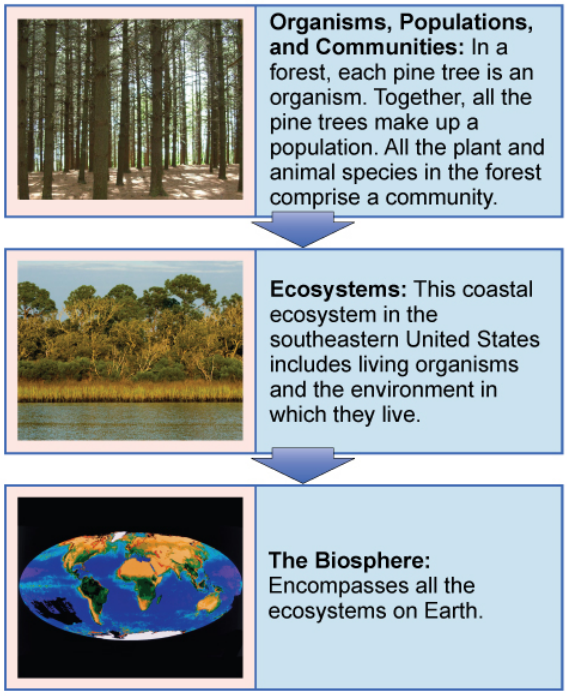


Figure 3.1.*a***:** Ecologists study within several biological levels of organization. (credit “organisms”: modification of work by "Crystl"/Flickr; credit “ecosystems”: modification of work by Tom Carlisle, US Fish and Wildlife Service Headquarters; credit “biosphere”: NASA)

Organismal Ecology

Researchers studying ecology at the organismal level are interested in the adaptations that enable individuals to live in specific habitats. These adaptations can be morphological, physiological, and behavioral. For instance, the Karner blue butterfly (*Lycaeides melissa samuelis;* figure 3.1.*b*) is considered a specialist because the females preferentially oviposit (that is, lay eggs) on wild lupine. This preferential adaptation means that the Karner blue butterfly is highly dependent on the presence of wild lupine plants for its continued survival.



Figure 3.1.*b*: The Karner blue butterfly (*Lycaeides melissa samuelis*) is a rare butterfly that lives only in open areas with few trees or shrubs, such as pine barrens and oak savannas. It can only lay its eggs on lupine plants. (credit: modification of work by J & K Hollingsworth, USFWS)

After hatching, the larval caterpillars emerge and spend four to six weeks feeding solely on wild lupine (figure 3.1.*c*). The caterpillars pupate (undergo metamorphosis) and emerge as butterflies after about four weeks. The adult butterflies feed on the nectar of flowers of wild lupine and other plant species. A researcher interested in studying Karner blue butterflies at the organismal level might, in addition to asking questions about egg laying, ask questions about the butterflies' preferred temperature (a physiological question) or the behavior of the caterpillars when they are at different larval stages (a behavioral question).



Figure 3.1.*c*: The wild lupine (*Lupinus perennis*) is the host plant for the Karner blue butterfly

Population Ecology

A population is a group of interbreeding organisms that are members of the same species living in the same area at the same time. (Organisms that are all members of the same species are called conspecifics.) A population is identified, in part, by where it lives, and its area of population may have natural or artificial boundaries: natural boundaries might be rivers, mountains, or deserts, while examples of artificial boundaries include mowed grass, manmade structures, or roads. The study of population ecology focuses on the number of individuals in an area and how and why population size changes over time. Population ecologists are particularly interested in counting the Karner blue butterfly, for example, because it is classified as federally endangered. However, the distribution and density of this species is highly influenced by the distribution and abundance of wild lupine. Researchers might ask questions about the factors leading to the decline of wild lupine and how these affect Karner blue butterflies. For example, ecologists know that wild lupine thrives in open areas where trees and shrubs are largely absent. In natural settings, intermittent wildfires regularly remove trees and shrubs, helping to maintain the open areas that wild lupine requires. Mathematical models can be used to understand how wildfire suppression by humans has led to the decline of this important plant for the Karner blue butterfly.

Community Ecology

A biological community consists of the different species within an area, typically a three-dimensional space, and the interactions within and among these species. Community ecologists are interested in the processes driving these interactions and their consequences. Questions about conspecific interactions often focus on competition among members of the same species for a limited resource. Ecologists also study interactions among various species; members of different species are called heterospecifics.

Examples of heterospecific interactions include predation, parasitism, herbivory, competition, and pollination. These interactions can have regulating effects on population sizes and can impact ecological and evolutionary processes affecting diversity.

For example, Karner blue butterfly larvae form mutualistic relationships with ants. Mutualism is a form of a long-term relationship that has coevolved between two species and from which each species benefits. For mutualism to exist between individual organisms, each species must receive some benefit from the other as a consequence of the relationship. Researchers have shown that there is an increase in the probability of survival when Karner blue butterfly larvae (caterpillars) are tended by ants. This might be because the larvae spend less time in each life stage when tended by ants, which provides an advantage for the larvae. Meanwhile, the Karner blue butterfly larvae secrete a carbohydrate-rich substance that is an important energy source for the ants. Both the Karner blue larvae and the ants benefit from their interaction.

Ecosystem Ecology

Ecosystem ecology is an extension of organismal, population, and community ecology. The ecosystem is composed of all the biotic components (living things) in an area along with the abiotic components (non-living things) of that area. Some of the abiotic components include air, water, and soil. Ecosystem biologists ask questions about how nutrients and energy are stored and how they move among organisms and the surrounding atmosphere, soil, and water.

The Karner blue butterflies and the wild lupine live in an oak-pine barren habitat. This habitat is characterized by natural disturbance and nutrient-poor soils that are low in nitrogen. The availability of nutrients is an important factor in the distribution of the plants that live in this habitat. Researchers interested in ecosystem ecology could ask questions about the importance of limited resources and the movement of resources, such as nutrients, though the biotic and abiotic portions of the ecosystem.

Career Connection: Ecologist

A career in ecology contributes to many facets of human society. Understanding ecological issues can help society meet the basic human needs of food, shelter, and health care. Ecologists can conduct their research in the laboratory and outside in natural environments (Figure 3.1.*d*). These natural environments can be as close to home as the stream running through your campus or as far away as the hydrothermal vents at the bottom of the Pacific Ocean. Ecologists manage natural resources such as white-tailed deer populations (*Odocoileus virginianus*) for hunting or aspen (*Populus* spp.) timber stands for paper production. Ecologists also work as educators who teach children and adults at various institutions including universities, high schools, museums, and nature centers. Ecologists may also work in advisory positions assisting local, state, and federal policymakers to develop laws that are ecologically sound, or they may develop those policies and legislation themselves. To become an ecologist requires an undergraduate degree, usually in a natural science. The undergraduate degree is often followed by specialized training or an advanced degree, depending on the area of ecology selected. Ecologists should also have a broad background in the physical sciences, as well as a sound foundation in mathematics and statistics.



Figure 3.1.*d***:** This landscape ecologist is releasing a black-footed ferret into its native habitat as part of a study. (credit: USFWS Mountain Prairie Region, NPS)

Attribution

Modified by Melissa Ha and Rachel Schleiger from Ecology and the Biosphere by OpenStax (licensed under CC-BY)

This page titled 3.1: Ecology Overview is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **4: Ecology Overview** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

3.2: Populations

Chapter Hook

Roadkill, we've all seen it on the roadside. It is just one very unfortunate outcomes of our extensive road networks. Roads, especially freeways and busy highways, also break apart habitats. When habitats are broken up into smaller pieces it weakens wildlife populations which need to move freely across landscapes for resources, refuge, and mates. Over long time scales, these fragmented habitats decrease species diversity, weaken species population gene pools, and can lead to extirpation (elimination of local population(s)) and extinction. As a result of numerous studies, ecologists have found one solution, wildlife crossings (aka green corridors or wildlife corridors). Wildlife crossings look like an underpass or overpass. However, they are filled with dirt and plants to look more natural. These crossings enable wildlife populations to move more freely, and safely, across landscapes all over the globe.



Figure 3.2.*a*: A newly constructed wildlife overpass over a Singapore highway. Image by Benjamin P. Y-H. Lee (licensed under CC-BY-4.0)

Populations are interacting and interbreeding groups of individuals from the same species in a common area. The study of **population ecology** focuses on population size and the factors that regulate population growth.

Attribution

Rachel Schleiger and Melissa Ha (CC-BY-NC)

3.2.1: Population Dispersion

3.2.2: Population Size

3.2.3: Population Growth and Regulation

3.2.4: Life History

3.2.5: Data Dive- Wildlife Corridors

3.2.6: Review

This page titled 3.2: Populations is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

3.2.1: Population Dispersion

The **dispersion pattern** (distribution pattern) of a population describes the arrangement of individuals within a habitat at a particular point in time, and broad categories of patterns are used to describe them. The three dispersion patterns are clumped, random, and uniform (figure 3.2.1.a).



Figure 3.2.1.a: Species may have a random, clumped, or uniform distribution. Plants such as (a) dandelions with wind-dispersed seeds tend to be randomly distributed (represented by dots in no particular pattern). Animals such as (b) elephants that travel in groups exhibit a clumped distribution (represented by dots clustered together). Territorial birds such as (c) penguins tend to have a uniform distribution (represented by dots evenly spaced from each other). (credit a: modification of work by Rosendahl; credit b: modification of work by Rebecca Wood; credit c: modification of work by Ben Tubby)

Individuals that are grouped into patches have a **clumped distribution**, or aggregated distribution. This can occur if resources are distributed unequally; for example, pipevine swallowtail caterpillars would be clumped in areas with their host plant, California pipevine. Clumped distribution may also reflect the locations of suitable microhabitats, such as an herbaceous (non-woody) plant species that only grows in the shade clustering under trees. Plants that drop their seeds straight to the ground, such as oak trees, may also have this distribution. Finally, social behavior in animals results in a clumped distribution, such as wolves hunting in a pack, a herd of elephants, or a school of fish traveling together for safety.

Populations that have a **random distribution** are not arranged in any particular pattern. Some individuals may be close together while others may be far apart. An example of random distribution occurs with dandelion and other plants that have wind-dispersed seeds that germinate wherever they happen to fall in favorable environments.

Individuals that are equally spaced in the environment have a **uniform distribution**. Saguaro cacti are evenly spaced due to limited resources in the desert. (There is not sufficient water to support two large cacti side-by-side.) Uniform distributions can result from **interference competition**, when individuals take pre-emptive measures to avoid competition for resources. For example, some plants that secrete substances inhibiting the growth of nearby individuals (such as the release of toxic chemicals by sage plants), a phenomenon called **allelopathy**. Another example of interference competition occurs in territorial animal species, such as penguins that maintain a defined territory for nesting. The territorial defensive behaviors of each individual create a regular pattern of distribution of similar-sized territories and individuals within those territories. Thus, the distribution of the individuals within a population provides more information about how they interact with each other than does a simple density measurement (see Population Size and Density).

Attribution

Modified by Melissa Ha from Population Demographics and Dynamics from *Environmental Biology* by Matthew R. Fisher (CC-BY)

This page titled 3.2.1: Population Dispersion is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **5.1: Population Dispersion** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.2.2: Population Size

Populations are dynamic entities. Their size and composition fluctuate in response to numerous factors, including seasonal and yearly changes in the environment, natural disasters such as forest fires and volcanic eruptions, and competition for resources between and within species. The study of populations is called **demography.**

Population Size and Density

Populations are characterized by their **population size** (total number of individuals) and their **population density** (number of individuals per unit area; figure 3.2.2.*a*). A population may have a large number of individuals that are distributed densely, or sparsely. There are also populations with small numbers of individuals that may be dense or very sparsely distributed in a local area. Population size can affect potential for adaptation because it affects the amount of genetic variation present in the population. Density can have effects on interactions within a population such as competition for food, the ability of individuals to find a mate, and diseases spread. (Dispersion patterns can also affect these factors; for example, a solitary species with a random distribution might have difficulty finding a mate when compared to social species clumped together in groups.) Smaller organisms tend to be more densely distributed than larger organisms (figure 3.2.2.*b*).

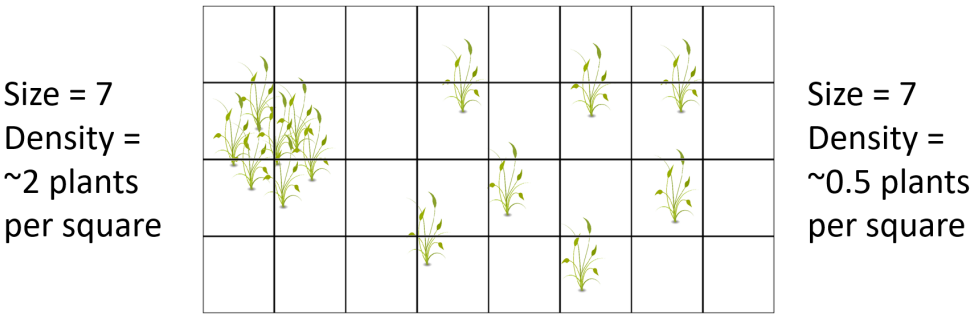


Figure 3.2.2.*a*: Both populations shown contain seven plants, so their populations size is the same. However, the population on the left has a higher population density (2 plants per square) than the population on the right (density = 0.5 plants per square). Image by Melissa Ha using Green Plant (public domain).

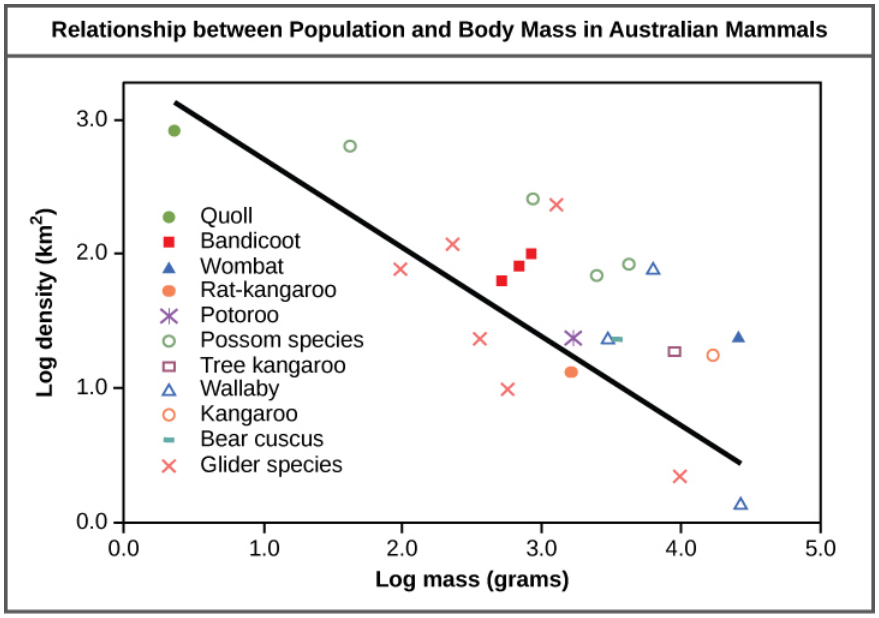


Figure 3.2.2.*b*: Australian mammals show a typical inverse relationship between population density and body size. Log mass in grams is on the x-axis, and log density in square kilometers (km^2) is on the y-axis. Each data point represents a different species, including the quoll, four bandicoot species, wombat, rat-kangaroo, potoroo, four possom species, tree kangaroo, three wallaby species, kangaroo, bear cuscus, and six glider species. As this graph shows, population density typically decreases with increasing body size. For example, quolls have the lowest body mass and the highest density. The heaviest wallaby species has the lowest population density.

Estimating Population Size

The most accurate way to determine population size is to count all of the individuals within the area. However, this method is usually not logistically or economically feasible, especially when studying large areas. Thus, scientists usually study populations by sampling a representative portion of each habitat and using this sample to make inferences about the population as a whole. The methods used to sample populations to determine their size and density are typically tailored to the characteristics of the organism being studied. For immobile organisms such as plants, or for very small and slow-moving organisms, a quadrat may be used. A **quadrat** is a square structure that is randomly located on the ground and used to count the number of individuals that lie within its boundaries (Figure 3.2.2.c). To obtain an accurate count using this method, the square must be placed at random locations within the habitat enough times to produce an accurate estimate.



Figure 3.2.2.c: High schoolers using a quadrat in an intertidal zone. Image by kqedquest (CC-BY-NC)

For smaller mobile organisms, such as mammals, a technique called **mark and recapture** is often used. This method involves marking captured animals and releasing them back into the environment to mix with the rest of the population. Later, a new sample is captured and scientists determine how many of the marked animals are in the new sample. This method assumes that the larger the population, the lower the percentage of marked organisms that will be recaptured since they will have mixed with more unmarked individuals. For example, if 80 field mice are captured, marked, and released into the forest, then a second trapping 100 field mice are captured and 20 of them are marked, the population size (N) can be determined using the following equation:

$$N = \frac{(\text{number marked first catch} \times \text{total number of second catch})}{\text{number marked second catch}} \tag{3.2.2.1}$$

Using our example, the equation would be:

$$\frac{(80 \times 100)}{20} = 400 \tag{3.2.2.2}$$

These results give us an estimate of 400 total individuals in the original population. The true number usually will be a bit different from this because of chance errors and possible bias caused by the sampling methods.

The mathematical methods required to estimate population sizes can be influenced by dispersion pattern.

Attribution

Modified by Melissa Ha from Population Demographics and Dynamics from *Environmental Biology* by Matthew R. Fisher (CC-BY)

This page titled 3.2.2: Population Size is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **5.2: Population Size** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.2.3: Population Growth and Regulation

Population ecologists make use of a variety of methods to model population dynamics. An accurate model should be able to describe the changes occurring in a population and predict future changes. The two simplest models of population growth use deterministic equations (equations that do not account for random events) to describe the rate of change in the size of a population over time. The first of these models, **exponential growth**, describes populations that increase in numbers without any limits to their growth. The second model, **logistic growth**, introduces limits to reproductive growth that become more intense as the population size increases. Neither model adequately describes natural populations, but they provide points of comparison.

The Population Growth Rate (*r*)

The **population growth rate** (sometimes called the rate of increase or per capita growth rate, *r*) equals the **birth rate** (*b*) minus the **death rate** (*d*) divided by the **initial population size** (N_0).

$$r = (b - d) \div N_0$$

Another method of calculating the population growth rate involves final and initial population size (figure 3.2.3.*a*). In this case, **population growth rate** (*r*) equals the **final population size** (N) minus the **initial population size** (N_0) and divided by the **initial population size** (N_0).

$$r = (N - N_0) \div N_0$$



Figure 3.2.3.*a*: How could we calculate the growth rate of this caribou population per year if there were 200 individuals in 2016 and 300 individuals in 2018? Image by K. Joly/NPS (public domain).

Doubling Time

The **doubling time** is how long it will take for a population to become twice its initial size. The **doubling time** (*t*) is equal to 0.69 divided by the **population growth rate** (*r*), written as a proportion.

$$t = 0.69 \div r$$

Population ecologists sometimes round this equation and calculate doubling time using the **"Rule of 70"** (dividing 70 by the population growth rate, written as a percentage). To express population growth rate as a percentage, it is multiplied by 100%. Thus, the 0.69 in the original doubling time equation is also multiplied by 100. This value (69) is rounded to 70 for simplicity.

Exponential Growth

Charles Darwin, in developing his theory of natural selection, was influenced by the English clergyman Thomas Malthus. Malthus published his book in 1798 stating that populations with abundant natural resources grow very rapidly. However, they limit further growth by depleting their resources. The early pattern of accelerating population size is called exponential growth (figure 3.2.3.*b*).

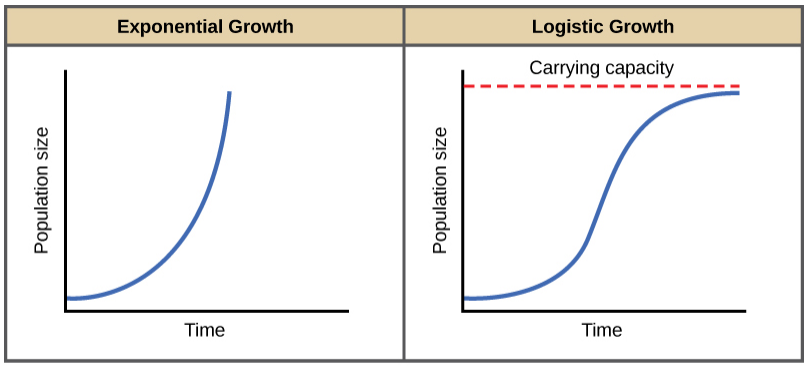


Figure 3.2.3.*b*: These graphs show exponential and logistic growth, with population size on the y-axis and time on the x-axis. When resources are unlimited, populations exhibit (a) exponential growth, shown in a J-shaped curve. When resources are limited, populations exhibit (b) logistic growth. In logistic growth, population expansion decreases as resources become scarce, and it levels off when the carrying capacity of the environment is reached. The logistic growth curve is S-shaped.

The best example of exponential growth in organisms is seen in bacteria. Bacteria are prokaryotes that reproduce quickly, about an hour for many species. If 1000 bacteria are placed in a large flask with an abundant supply of nutrients (so the nutrients will not become quickly depleted), the number of bacteria will have doubled from 1000 to 2000 after just an hour (figure 3.2.3.*c*). In another hour, each of the 2000 bacteria will divide, producing 4000 bacteria. After the third hour, there should be 8000 bacteria in the flask. The important concept of exponential growth is that the growth rate—the number of organisms added in each reproductive generation—is itself increasing; that is, the population size is increasing at a greater and greater rate. After 24 of these cycles, the population would have increased from 1000 to more than 16 billion bacteria. When the population size, *N*, is plotted over time, a J-shaped growth curve is produced (figure 3.2.3.*b*).



Figure 3.2.3.*c*: Bacteria growing with abundant nutrients in a flask will exhibit exponential growth. Image by Soledad Mirand-Rottman (CC-BY-SA).

The bacteria-in-a-flask example is not truly representative of the real world where resources are usually limited. However, when a species is introduced into a new habitat that it finds suitable, it may show exponential growth for a while. In the case of the bacteria in the flask, some bacteria will die during the experiment and thus not reproduce; therefore, the growth rate is lowered from a maximal rate in which there is no mortality.

Logistic Growth

Extended exponential growth is possible only when infinite natural resources are available; this is not the case in the real world. Charles Darwin recognized this fact in his description of the “struggle for existence,” which states that individuals will compete,

with members of their own or other species, for limited resources. The successful ones are more likely to survive and pass on the traits that made them successful to the next generation at a greater rate (natural selection). To model the reality of limited resources, population ecologists developed the logistic growth model.

In the real world, with its limited resources, exponential growth cannot continue indefinitely. Exponential growth may occur in environments where there are few individuals and plentiful resources, but when the number of individuals gets large enough, resources will be depleted and the growth rate will slow down. Eventually, the growth rate will plateau or level off (figure 3.2.3.*b*). This population size, which is determined by the maximum population size that a particular environment can sustain, is called the **carrying capacity**, symbolized as *K*. In real populations, a growing population often overshoots its carrying capacity and the death rate increases beyond the birth rate causing the population size to decline back to the carrying capacity or below it. Most populations usually fluctuate around the carrying capacity in an undulating fashion rather than existing right at it.

A graph of logistic growth yields the S-shaped curve (figure 3.2.3.*b*). It is a more realistic model of population growth than exponential growth. There are three different sections to an S-shaped curve. Initially, growth is exponential because there are few individuals and ample resources available. Then, as resources begin to become limited, the growth rate decreases. Finally, the growth rate levels off at the carrying capacity of the environment, with little change in population number over time.

While bacteria in a flask with abundant nutrients might initially exhibit exponential growth, bacteria grown with limited nutrients can exhibit logistic growth (figure 3.2.3.*d*).

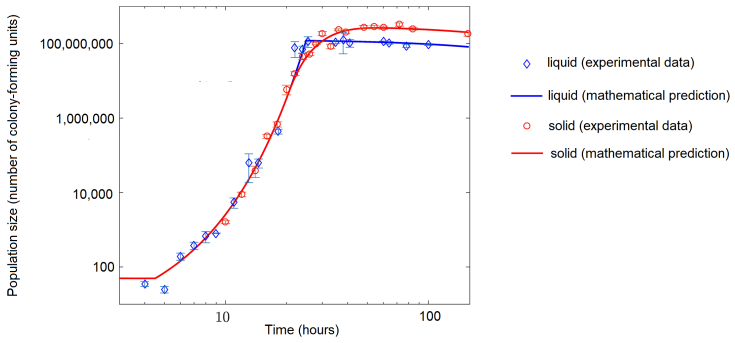


Figure 3.2.3.*d*: The growth of the bacterium *Escherichia coli* (*E. coli*) with limited nutrients over time in hours. Population size is in number of colony-forming units, which is a cell capable of dividing to produce a colony of cells. The bacteria were grown in liquid (blue diamonds) or on a semisolid medium (agar, red circles). The points represent actual measurements of populations size, and the lines represent a mathematical model, basically a prediction that is based on the data points. The data show logistic population growth, particularly in solid media. Initially, population size increases rapidly (concave part of curve), but the growth rate then stabilizes (straight line) before decreasing (convex part of curve) and leveling near the carrying capacity. Image modified from Shao X, Mugler A, Kim J, Jeong HJ, Levin BR, Nemenman I (2017) Growth of bacteria in 3-d colonies. PLoS Comput Biol 13(7): e1005679 (CC-BY).

In some populations, there are variations to the S-shaped curve. Examples in wild populations include sheep and harbor seals (figure 3.2.3.*e*). In both examples, the population size exceeds the carrying capacity for short periods of time and then falls below the carrying capacity afterwards. This fluctuation in population size continues to occur as the population oscillates around its carrying capacity. Still, even with this oscillation the logistic model is confirmed.

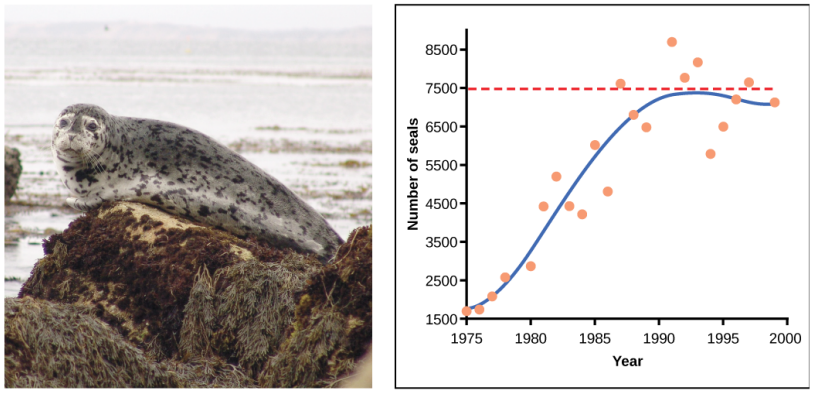


Figure 3.2.3.*e*: A natural population of seals shows logistic growth. The number of seals, ranging from 1500 to 8500, is on the y-axis, and the year, ranging from 1975-2000, is on the x-axis. Orange dots represent the observed population size each year, and the blue line represent the trend. From 1975 to 1983, the blue line curves upwards, representing an increase in both population size and population growth rate. From 1983 to 1992, the curve flattens, representing that population size is still increasing, but population growth rate is decreasing. After 1992, population fluctuates near the carrying capacity (dotted, horizontal line).

The logistic population growth model is not the only way that populations respond to limited resources. In some populations, growth is exponential until resources run low, wastes accumulate, or disease spreads (see limiting factors below), and the population then crashes. Thus, population growth rate (and size) may plummet rapidly instead of tapering as it approaches the carrying capacity.

Population Dynamics and Regulation

The logistic model of population growth, while valid in many natural populations and a useful model, is a simplification of real-world population dynamics. Implicit in the model is that the carrying capacity of the environment does not change, which is not the case. The carrying capacity varies annually. For example, some summers are hot and dry whereas others are cold and wet; in many areas, the carrying capacity during the winter is much lower than it is during the summer. Furthermore, some factors (**growth factors**) increase population growth rate while other factors (**limiting factors**) slow population growth. Examples of growth factors are resources like food, water, and space. Limiting factors can be classified as density-dependent or density-independent.

Density-dependent Regulation

Most **density-dependent factors** are biological in nature (biotic). Usually, the denser a population is, the greater its mortality rate. An example of density-dependent regulation is shown in figure 3.2.3.*f* with results from a study focusing on the giant intestinal roundworm (*Ascaris lumbricoides*), a parasite of humans and other mammals. Denser populations of the parasite exhibited lower fecundity: they contained fewer eggs. One possible explanation for this is that females would be smaller in more dense populations (due to limited resources) and that smaller females would have fewer eggs. This hypothesis was tested and disproved in a 2009 study which showed that female weight had no influence. The actual cause of the density-dependence of fecundity in this organism is still unclear and awaiting further investigation.

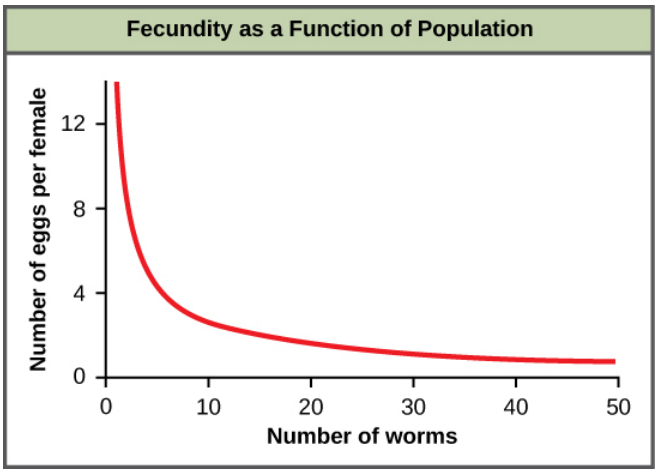


Figure 3.2.3.f: In this population of roundworms, fecundity (number of eggs per female) decreases with population size (number of worms). (Note that population size rather than population density is shown here, but population density does correlate with population size.) The number of eggs decreases rapidly at first, then levels off between 30 to 50 worms.

Density-dependent factors include predation, parasitism, herbivory, competition, and accumulation of waste. As a population increases, its predators are able to harvest it more easily. Prey density also affects population growth rate of predators: low prey density increases the mortality of its predator because it has more difficulty locating its food sources.

Parasites are able to pass from host to host more easily as the population density of the host increases. For this reason, epidemics among humans are particularly severe in cities. In fact, for most of the period since humans began living in cities, city populations have been maintained only through continual immigration from the countryside. Not until the development of community sanitation, immunization, and other public health measures did cities avoid periodic sharp drops in population as a result of epidemics. The recurrent epidemics of the "black death" in Europe that began in the fourteenth century caused a sharp decline in population. In just three years (1348–1350), at least one-quarter of the population of Europe died from the disease (probably plague).

Similarly, herbivores can more easily spread between individual plants in a dense population. This is why strip cropping (see Sustainable Agriculture) helps control pests. An herbivore or plant pathogen may infect one row of plants, but it is less likely to spread to more distant rows of that species.

While **interspecific competition** occurs between different species, **intraspecific competition** occurs when members of the same species harm each other by using the same resources. For example, in the summer of 1980, much of southern New England was struck by an infestation of the gypsy moth (figure 3.2.3.g). As the summer wore on, the larvae (caterpillars) pupated, the hatched adults mated, the females laid masses of eggs (each mass containing several hundred eggs) on virtually every tree in the region. In early May of 1981, the young caterpillars that hatched from these eggs began feeding and molting.



Figure 3.2.3.g: Gypsy moth caterpillars face intraspecific competition at high densities. Image by Editor at Large (CC-BY-SA).

The results were dramatic: In 72 hours, a 50-ft beech tree or a 25-ft white pine tree would be completely defoliated. Large patches of forest began to take on a winter appearance with their skeletons of bare branches. In fact, the infestation was so heavy that many trees were completely defoliated before the caterpillars could complete their larval development. The result: a massive die-off of the animals; very few succeeded in completing metamorphosis. Here, then, was a dramatic example of how competition among members of one species for a finite resource - in this case, food - caused a sharp drop in population. The effect was clearly density-dependent. The lower population densities of the previous summer had permitted most of the animals to complete their life cycle.

Density-independent Regulation

Density-independent factors, typically physical or chemical in nature (abiotic), influence the mortality of a population regardless of its density, including weather (figure 3.2.3.h), natural disasters (earthquakes, volcanoes, fires, etc.), and pollution. An individual deer may be killed in a forest fire regardless of how many deer happen to be in that area. Its chances of survival are the same whether the population density is high or low. The same holds true for cold winter weather.



Figure 3.2.3.h: Severe weather, like Hurricane Florence of 2018, can reduce population size regardless of density. Image by NASA Goddard Space Flight Center (CC-BY)

In real-life situations, population regulation is very complicated and density-dependent and independent factors can interact. A dense population that is reduced in a density-independent manner by some environmental factor(s) will be able to recover differently than a sparse population. For example, a population of deer affected by a harsh winter will recover faster if there are more deer remaining to reproduce.

Attributions

Modified by Melissa Ha from the following sources:

- Population Growth and Regulation from *Environmental Biology* by Matthew R. Fisher (CC-BY)
- Principles of Population Growth and The Human Population from *Biology* by John W. Kimball (CC-BY)
- Population Dynamics and Regulation from *General Biology* by OpenStax (CC-BY)

This page titled 3.2.3: Population Growth and Regulation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **5.3: Population Growth and Regulation** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

3.2.4: Life History

Life history characteristics, are those that relate to the survival, mortality, and reproduction of a species. Examples include birth rates, age at first reproduction, the number of offspring, and death rates. These characteristics evolve just like physical traits or behavior, leading to adaptations that affect population growth.

Reproductive Strategies

Population ecologists have hypothesized that suites of characteristics may evolve in species that lead to particular adaptations to their environments. These adaptations impact the kind of population growth their species experience. Population ecologists have described a continuum of life-history "strategies" with *K*-selected species on one end and *r*-selected species on the other (table 3.2.4.*a*). ***K*-selected species** are adapted to stable, predictable environments. Populations of *K*-selected species tend to exist close to their carrying capacity (which is represented by the letter "K" in the equation for logistic population growth). These species tend to have larger, but fewer, offspring, contribute large amounts of resources to each offspring, and have long generation times. Elephants would be an example of a *K*-selected species (figure 3.2.4.*a*). When a habitat becomes filled with a diverse collection of creatures competing with one another for the necessities of life, the advantage shifts to *K*-strategists. *K*-strategists have stable populations that are close to *K*. There is nothing to be gained from a high growth rate (*r*). The species will benefit most by a close adaptation to the conditions of its environment.

Table 3.2.4.*a***:** Comparison of *K*- and *r*-selected species

Characteristics of *K*-selected species	Characteristics of *r*-selected species
Mature late	Mature early
Greater longevity	Lower longevity
Increased parental care	Decreased parental care
Strong competitors	Weak competitors (strong colonizers)
Fewer offspring	More offspring
Larger offspring	Smaller offspring



(a) K-selected species



(b) r-selected species

Figure 3.2.4.a: (a) Elephants are considered K-selected species as they live long, mature late, and provide long-term parental care to few offspring. Oak trees produce many offspring that do not receive parental care, but are considered K-selected species based on longevity and late maturation. (b) Dandelions and jellyfish are both considered r-selected species as they mature early, have short lifespans, and produce many offspring that receive no parental care.

***r*-selected species** are adapted to unstable and unpredictable environments. They have large numbers of small offspring. Animals that are *r*-selected do not provide a lot of resources or parental care to offspring, and the offspring are relatively self-sufficient at birth. Examples of *r*-selected species are marine invertebrates such as jellyfish and plants such as the dandelion (figure 3.2.4.a). *r*-strategists have short life spans and reproduce quickly, resulting in short generation times. For example, the housefly can produce 7 generations each year (each of about 120 young). Ragweed is well-adapted to exploiting its environment in a hurry - before competitors can become established. It grows rapidly and produces a huge number of seeds (after releasing its pollen, the bane of many hay fever sufferers). Because ragweed's approach to continued survival is through rapid reproduction (a high value of *r)* it is called an *r*-strategist. Other weeds, many insects, and many rodents are also *r*-strategists. If fact, if we consider an organism a pest, it is probably an *r*-strategist.

The two extreme strategies are at two ends of a continuum on which real species life histories will exist. In addition, life history strategies do not need to evolve as suites but can evolve independently of each other; therefore, each species may have some characteristics that trend toward one extreme or the other. Nevertheless, the *r*- and *K*-selection theory provides a foundation for a more accurate life history framework. New demographic-based models of life history evolution have been developed, which incorporate many ecological concepts included in *r*- and *K*-selection theory as well as population age structure and mortality factors.

Life Tables

Life tables provide important information about the life history of an organism and the life expectancy of individuals at each age. They are modeled after actuarial tables used by the insurance industry for estimating human life expectancy. Life tables may include the probability of each age group dying before their next birthday, the percentage of surviving individuals dying at a particular age interval (their mortality rate, and their life expectancy at each interval). An example of a life table is shown in Table 3.2.4.*b* from a study of Dall mountain sheep, a species native to northwestern North America. Notice that the population is divided into age intervals (column A).

As can be seen from the mortality rate data (column D), a high death rate occurred when the sheep were between six months and a year old, and then increased even more from 8 to 12 years old, after which there were few survivors. The data indicate that if a sheep in this population were to survive to age one, it could be expected to live another 7.7 years on average, as shown by the life-expectancy numbers in column E. Life tables often include more information than shown in Table 3.2.4.*b*, such as fecundity (reproduction) rates for each age group.

Table 3.2.4.*b***:** Life Table of Dall Mountain Sheep

Age interval (years)	Number dying in age interval out of 1000 born	Number surviving at beginning of age interval out of 1000 born	Mortality rate per 1000 alive at beginning of age interval	Life expectancy or mean lifetime remaining to those attaining age interval
0–0.5	54	1000	54.0	7.06
0.5–1	145	946	153.3	—
1–2	12	801	15.0	7.7
2–3	13	789	16.5	6.8
3–4	12	776	15.5	5.9
4–5	30	764	39.3	5.0
5–6	46	734	62.7	4.2
6–7	48	688	69.8	3.4
7–8	69	640	107.8	2.6
8–9	132	571	231.2	1.9
9–10	187	439	426.0	1.3
10–11	156	252	619.0	0.9
11–12	90	96	937.5	0.6
12–13	3	6	500.0	1.2
13–14	3	3	1000	0.7

Survivorship Curves

Another tool used by population ecologists is a **survivorship curve**, which is a graph of the number of individuals surviving at each age interval versus time. These curves allow us to compare the life histories of different populations (figure 3.2.4.*c*). There are three types of survivorship curves. In a **type I survivorship curve**, mortality is low in the early and middle years and occurs mostly in older individuals. Organisms exhibiting a type I survivorship typically produce few offspring and provide good care to the offspring increasing the likelihood of their survival. Humans and most mammals exhibit a type I survivorship curve. In **type II survivorship curves**, mortality is relatively constant throughout the entire life span, and mortality is equally likely to occur at any point in the life span. Many bird populations provide examples of an intermediate or type II survivorship curve. In **type III**

survivorship curves, early ages experience the highest mortality with much lower mortality rates for organisms that make it to advanced years. Type III organisms typically produce large numbers of offspring, but provide very little or no care for them. Trees and marine invertebrates exhibit a type III survivorship curve because very few of these organisms survive their younger years, but those that do make it to an old age are more likely to survive for a relatively long period of time.

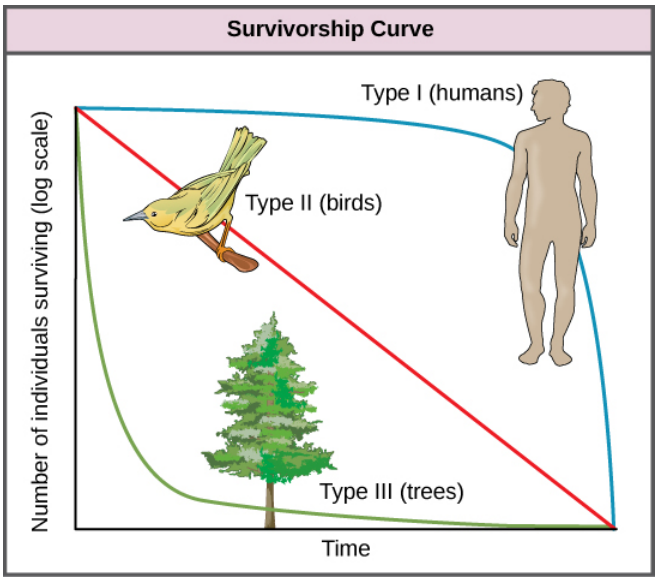


Figure 3.2.4.*c*: Survivorship curves show the chance of survival in a population according to age. The number of individuals surviving (on a log scale) is on the y-axis, and time (representing relative age) is on the x-axis. Humans and most mammals have a Type I survivorship curve, because death primarily occurs in the older years. The Type I survivorship curve is represented by a convex curve, which rapidly declines at the end. Birds have a Type II survivorship curve, as death at any age is equally probable. The Type II survivorship curve is represented by a straight line, going from the upper left to lower right of the graph. Trees have a Type III survivorship curve because very few survive the younger years, but after a certain age, individuals are much more likely to survive. The Type III survivorship curve is represented by a concave curve, which initially declines rapidly and then flattens.

While is no exact association between reproductive strategies (*K*- or *r*-selected) and survivorship curves (Type I, II, or III), *K*-selected species are more likely to have a Type III survivorship curve. *r*-selected species tend to have a Type I survivorship curve.

Attribution

Modified by Melissa Ha from the following sources:

- Population Dynamics and Regulation from *General Biology* by OpenStax (CC-BY)
- Population Demographics and Dynamics from *Environmental Biology* by Matthew R. Fisher (CC-BY)
- Principles of Population Growth from *Biology* by John W. Kimball (CC-BY)

This page titled 3.2.4: Life History is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **5.4: Life History** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.2.5: Data Dive- Wildlife Corridors

Overview

In 1999 the Sonoran Desert Conservation Plan was created to help wildlife populations and habitats. As a result of this plan, biologists were able to identify landscapes with wildlife barriers and get funding to install wildlife corridors. On State Route (SR) 77 two wildlife corridors were installed, one overpass and one underpass. These corridors were the first of their kind to be built in the Sonoran Desert so it was difficult to estimate what level of success they would have. As such, a monitoring plan was put in place to document which species used the corridors and how often they used it as time passed since their installation. The figure below illustrates the usage of the wildlife corridors 2+ years after their installation:

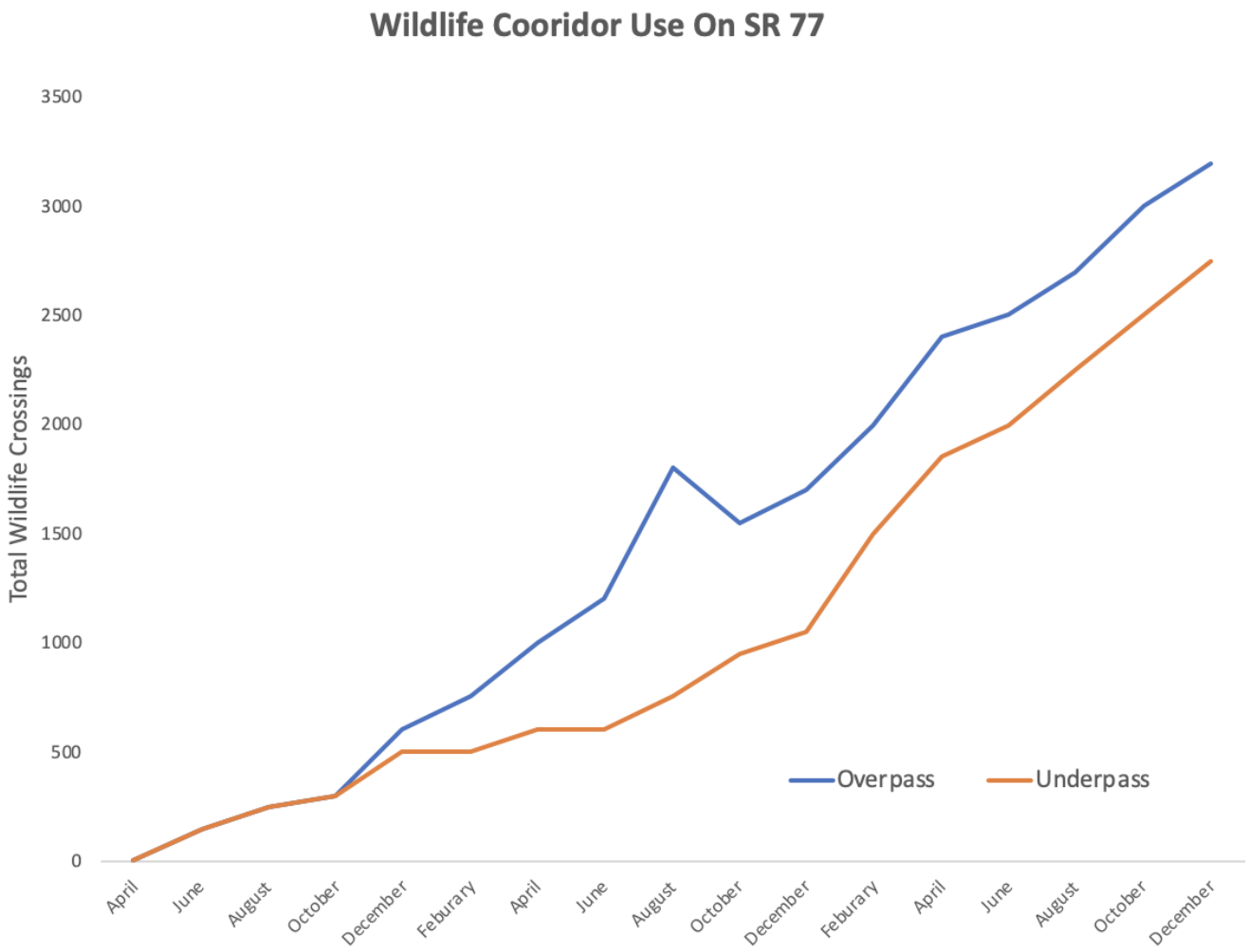


Figure 3.2.5.*a*: Total wildlife crossings after installation for overpass and underpass on SR 77. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Santa Catalina - Tortolita Mountain Corridor Progress Report

Questions

1. What is the independent (explanatory) variable and the dependent (response) variable?
2. What question(s) are the authors trying to answer with this graph?
3. Does this graph show a positive or negative relationship between the variables? Support your answer by referring to trends observed in the graph.
4. Do you think like the authors are satisfied with the results in the graph? Why?

5. How can we use the results of this graph to inform future wildlife corridor installations?
6. What information/patterns is not clear from this graph?

Raw Data For Above Graph(s)

Table 3.2.5.*a***:** Raw data for total wildlife crossings after installation for overpass and underpass on SR 77. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Santa Catalina - Tortolita Mountain Corridor Progress Report

Month	Overpass	Underpass
April	0	0
June	150	150
August	249	250
October	296	294
December	600	500
February	750	502
April	1000	600
June	1200	605
August	1800	750
October	1550	950
December	1700	1050
February	2000	1500
April	2400	1850
June	2500	2000
August	2700	2250
October	3000	2500
December	3200	2750

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 3.2.5: Data Dive- Wildlife Corridors is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **5.5: Data Dive- Wildlife Corridors** by Melissa Ha and Rachel Schleiger has no license indicated.

3.2.6: Review

Summary

After completing this chapter you should be able to...

- Describe three different population dispersion patterns.
- Differentiate between population size and density.
- Explain how ecologists measure population size.
- Calculate population growth rate (*r*) and doubling time (*t*).
- Distinguish between exponential and logistic population growth models, explaining the role of the carrying capacity in logistic growth.
- Provide examples of density-dependent and density-independent factors that regulate populations.
- Compare *K*-selected and *r*-selected reproductive strategies.
- Interpret life tables.
- Compare type I, type II, and type III survivorship curves.

Populations are interacting, interbreeding groups of individuals from the same species. Ecologists measure characteristics of populations: **dispersion pattern**, **population size**, and **population density**. Populations with unlimited resources grow **exponentially**—with an accelerating growth rate. When resources become limiting, populations follow a **logistic growth curve** in which population size will level off at the **carrying capacity**. **Density-dependent factors** limit population growth as they reach their carrying capacity and include biotic factors such as predation, competition, and disease. Density-independent factors, such as storms and fires, are abiotic and decrease population size regardless of density.

Several frameworks explain how **life history** can influence population dynamics. ***K*-selected species** tend to have long life spans and produce few offspring with much parental care whereas ***r*-selected species** mature and reproduce rapidly, producing many offspring and offering little parental care. **Life tables** are useful to calculate life expectancies of individual population members. **Survivorship curves** show the number of individuals surviving at each age interval plotted versus time.

Attribution

Modified by Melissa Ha from Community and Population Ecology and Chapter Resources from *Environmental Biology* by Matthew R. Fisher (CC-BY)

This page titled 3.2.6: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **5.6: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

3.3: Communities

Chapter Hook

Until the late 1800s, Bison (*Bison bison*) numbered tens of millions in the Great Plains of the United States. By 1890, roughly 1000 Bison were left because the United States government campaign to eradicate the native people, their culture, and habitats they relied on. Slowly, a movement started to try and save Bison from extinction. It took until the early 2000's for Bison numbers to reach a half a million. During this time, scientists were able to observe the Bison reintroduction back into the Great Plains. Bison were found to be the most critical species to restoring and maintaining the function and diversity in the Great Plains community. Both plants and animal populations in the community were strengthened from the return of the Bison. Understanding community dynamics is essential to conserving and restoring these systems and the species that define them. This is particularly critical for communities with one particular species that acts as a keystone to the health of the system.



Figure 3.3.*a***:** American bison with starlings on its back. Image by NPS photos/Kim Acker (Public Domain)

Populations typically do not live in isolation from other species. Populations that interact within a given area form a **community**. The organisms that form a community are found in **habitats**, physical environments where organisms live; however, **biotic** (living) components are considered part of a community. Scientists study ecology at the community level to understand how species interact with each other and compete for the same resources.

Attribution

Modified by Rachel Schleiger and Melissa Ha from Community Ecology from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

3.3.1: Biotic Interactions

- 3.3.1.1: Trophic Interactions
  - 3.3.1.1.1: Predation
  - 3.3.1.1.2: Herbivory
  - 3.3.1.1.3: Parasitism
  - 3.3.1.1.4: Food Chains and Food Webs

3.3.1.2: Competition
3.3.1.3: Facilitation
3.3.1.4: Symbiosis

3.3.2: Community Structure and Dynamics

3.3.3: Data Dive- Bison Impacts In Prairies

3.3.4: Review

This page titled 3.3: Communities is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

SECTION OVERVIEW

3.3.1: Biotic Interactions

Biotic interactions refer to the relationships among organisms. They can be **intraspecific** (between members of the same species) or **interspecific** (between members of different species). When at least one of the interactants is harmed, the relationship is called an **antagonism. Trophic interactions**, in which one species consumes another, are antagonisms. **Competition** is another antagonism in which species at the same trophic level (that eat the same things) interact through using the same resources. Interactions in which at least one species benefits and neither is harmed are called **facilitation**, which can be categorized as **commensalism** or **mutualism.**

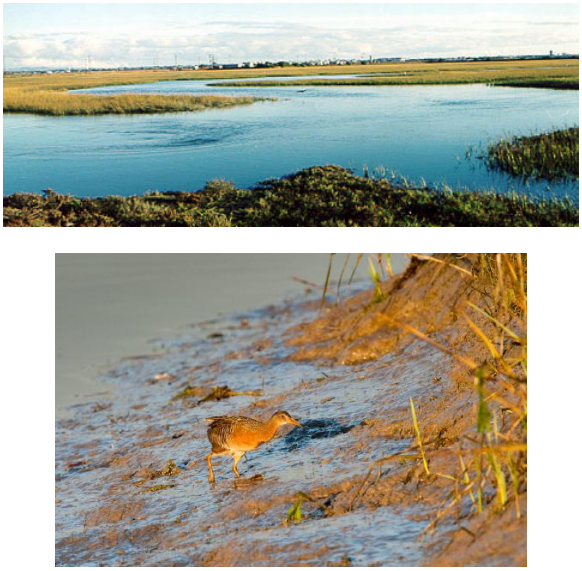


Figure 3.3.1.*a*: A variety of biotic interactions occur among the species at The Seal Beach National Wildlife Refuge (top). For example, the Light-Footed Clapper Rail (bottom) benefits from the cordgrass in which it nests. However, predators such as red foxes harm this endangered bird. The refuge manages this harmful interaction by installing fencing to exclude the red fox, which historically did not occur in the area. Top image and bottom image by US Fish & Wildlife Service (public domain).

Attribution

Modified by Melissa Ha from Community Ecology from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

Biotic interactions

3.3.1.1: Trophic Interactions

3.3.1.1.1: Predation

3.3.1.1.2: Herbivory

3.3.1.1.3: Parasitism

3.3.1.1.4: Food Chains and Food Webs

3.3.1.2: Competition

3.3.1.3: Facilitation

3.3.1.4: Symbiosis

This page titled 3.3.1: Biotic Interactions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

SECTION OVERVIEW

3.3.1.1: Trophic Interactions

Trophic interactions occur when one organism feeds on another. The three main types of trophic interactions are **predation** (figure 3.3.1.1.*a*), **herbivory**, or **parasitism**. During these interactions, one species benefits by gaining food at the expense of the other, which either dies or loses nutrients, tissues, or organs (such as leaves). Trophic interactions involve the flow of energy, and the trophic interactions in a community can be represented by **food chains** and **food webs**.



Figure 3.3.1.1.*a*: A hawk (*Buteo jamaicensis*) eats a vole (*Microtus californicus*), exemplifying predation, a common trophic interaction. Image by Jrockley (public domain).

Attribution

Melissa Ha (CC-BY-NC)

Trophic Interactions

3.3.1.1.1: Predation

3.3.1.1.2: Herbivory

3.3.1.1.3: Parasitism

3.3.1.1.4: Food Chains and Food Webs

This page titled 3.3.1.1: Trophic Interactions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

3.3.1.1.1: Predation

Perhaps the classical example of a biotic interaction is the predator-prey relationship. **Predation** occurs when one species (the **predator**) kills and eats multiple **prey** over its lifetime.

Predator-Prey Dynamics

Population sizes of predators and prey in a community are not constant over time, and they may vary in cycles that appear to be related. For example, 100 years of trapping data from North America shows how population dynamics of the lynx (predator) and the snowshoe hare (prey) cycle (figure 3.3.1.1.1.a-b).



Figure 3.3.1.1.1.a: The Canada lynx is an example of a predator. Image by Michael Zahra (CC-BY-SA).



Figure 3.3.1.1.1.b: The snowshoe hare is the prey of the Canada lynx. Image by the National Park Service (public domain).

This cycling of predator and prey population sizes has a period of approximately ten years, with the predator population lagging one to two years behind the prey population (figure 3.3.1.1.1.a-c). An apparent explanation for this pattern is that as the hare numbers increase, there is more food available for the lynx, allowing the lynx population to increase as well. When the lynx population grows to a threshold level, however, they kill so many hares that hare numbers begin to decline, followed by a decline in the lynx population because of scarcity of food. When the lynx population is low, the hare population size begins to increase due, in part, to low predation pressure, starting the cycle anew.

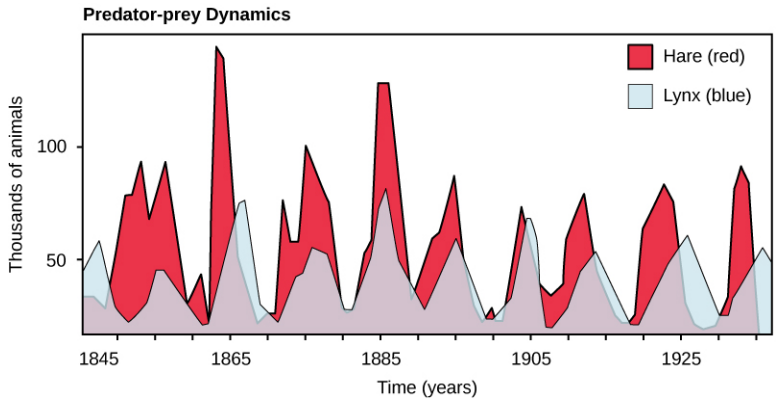


Figure 3.3.1.1.1.*c*: The cycling of Canada lynx and snowshoe hare populations in Northern Ontario is an example of predator-prey dynamics. The x-axis shows time in years, and the y-axis shows population size in thousands of animals. The lynx population is represented in blue (shorter peaks), and the hare population is represented in red (taller peaks). Both populations increase and decrease in 10-year cycles, with changes in the lynx population lagging a few years behind that of hare.

The dynamic of closely matched population cycling occurs because the snowshoe hare is the primary food source for the lynx. When predators have generalists, feeding on a variety of prey species, and prey are consumed by a variety of predators species, their population dynamics are less likely to cycle together.

Adaptations of Predators

Predators have a variety of adaptations to catch and consume prey, and the specific adaptations depend on the predator. For example, raptors, such as owls and hawks, have hooked beaks for tearing flesh and talons for grabbing prey. Mammal predators often have sharp teeth and claws. Some predators can run quickly to chase prey, and others "sit and wait", lunging forward when prey pass them. Some predators, like rattlesnakes and tarantulas, subdue their prey by injecting them with venom (figure 3.3.1.1.1.*d*). Predators often have large eyes located forward (like a wolf) rather than eyes spaced far apart on the sides of the head (like a sheep). Forward-facing eyes allow for depth perception, which is key to tracking prey. In contrast, peripheral vision is expanded when eyes are located to the sides, and this helps prey identify threats.



Figure 3.3.1.1.1.*d*: This Asian birdeater tarantula uses its fangs to inject its prey with poisonous venom. Image by Matt Reinbold (CC-BY-SA).

Defensive Traits in Prey

Prey evolve mechanical, chemical, physical, or behavioral defenses against predators. Some prey have armor (such as a turtle shell or the bony plates protecting armadillos), a **mechanical defense** that reduces predation by discouraging physical contact. Many animals produce or obtain **chemical defenses** from plants and store them to prevent predation. Other species use their body shape and coloration as **camouflage** to avoid being detected by predators, an example of a **physical defense**. The tropical walking stick is an insect with the coloration and body shape of a twig, which makes it very hard to see when it is stationary against a background

of real twigs (figure 3.3.1.1.1.*e*-a). In another example, the chameleon can change its color to match its surroundings (figure 3.3.1.1.1.*e*-b).



(a) (b)

Figure 3.3.1.1.1.*e*: (a) The tropical walking stick and (b) the chameleon use their body shape and/or coloration to prevent detection by predators. (credit a: modification of work by Linda Tanner; credit b: modification of work by Frank Vassen)

Some species use coloration as a way of warning predators that they are distasteful or poisonous. For example, the monarch butterfly caterpillar sequesters poisons from its food (plants and milkweeds) to make itself poisonous or distasteful to potential predators. The caterpillar is bright yellow and black to advertise its toxicity. The caterpillar is also able to pass the sequestered toxins on to the adult monarch, which is also dramatically colored black and red as a warning to potential predators. Fire-bellied toads produce toxins that make them distasteful to their potential predators (figure 3.3.1.1.1.*f*). They have bright red or orange coloration on their bellies, which they display to a potential predator to advertise their poisonous nature and discourage an attack. Warning coloration only works if a predator uses eyesight to locate prey and can learn—a naïve predator must experience the negative consequences of eating one before it will avoid other similarly colored individuals.



Figure 3.3.1.1.1.*f*: The fire-bellied toad has bright coloration on its belly that serves to warn potential predators that it is toxic. (credit: modification of work by Roberto Verzo)

While some predators learn to avoid eating certain potential prey because of their coloration, other species have evolved mechanisms to mimic this coloration to avoid being eaten, even though they themselves may not be unpleasant to eat or contain toxic chemicals. In some cases of **mimicry**, a harmless species imitates the warning coloration of a harmful species. Assuming they share the same predators, this coloration then protects the harmless ones. Many insect species mimic the coloration of wasps, which are stinging, venomous insects, thereby discouraging predation (figure 3.3.1.1.1.*g*). In other cases of mimicry, multiple species share the same warning coloration, but all of them actually have defenses. The commonness of the signal improves the compliance of all the potential predators.



(a) (b)

Figure 3.3.1.1.1.*g*: One form of mimicry is when a harmless species mimics the coloration of a harmful species, as is seen with the (a) wasp (*Polistes* sp.) and the (b) hoverfly (*Syrphus* sp.). (credit: modification of work by Tom Ings)

Running from predators, hiding, and playing dead are examples of **behavioral defenses**. Some prey also exhibit behaviors that threaten predators. For example, snapping turtles stretch their legs to appear larger and snap aggressively at predators.

Attribution

Modified by Melissa Ha from Community Ecology from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 3.3.1.1.1: Predation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **6.1.1.1: Predation** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

3.3.1.1.2: Herbivory

When an animal feed on parts of another organism (usually a plant), **herbivory** occurs. For example, the koala eats the eucalyptus leaves (figure 3.3.1.1.2.a) or a cicada feeds a plant's sap (figure 3.3.1.1.2.b). Like predators, herbivores may feed on multiple individuals over their lives; however, herbivores do not necessarily kill the plants that they eat.



Figure 3.3.1.1.2.a: Koalas are herbivores that feed exclusively on eucalyptus leaves. They are thus considered specialists because the specialize on just one type of food. Image by Dr. Umesh Behari Mathur (CC-BY).



Figure 3.3.1.1.2.b: This floury baker cicada feeds on plant sap by piercing the tree with needle-like mouth parts. Image by Toby Hudson (CC-BY-SA).

Plant species have evolved numerous defenses that reduce herbivory. Thorns are an example of a mechanical defense that discourages large herbivores from feeding on a plant (Figure 3.3.1.1.2.c-a). Many plant species produce secondary plant compounds that serve no function for the plant except that they are toxic to animals and discourage consumption, acting as a chemical defense. For example, the foxglove produces several compounds, including digitalis, that are extremely toxic when eaten (Figure 3.3.1.1.2.c-b). (Biomedical scientists have repurposed the chemical produced by foxglove as a heart medication, which has saved lives for many decades.)



(a) (b)

Figure 3.3.1.1.2.*c*: The (a) honey locust tree uses thorns, a mechanical defense, against herbivores, while the (b) foxglove uses a chemical defense: toxins produces by the plant can cause nausea, vomiting, hallucinations, convulsions, or death when consumed. (credit a: modification of work by Huw Williams; credit b: modification of work by Philip Jägenstedt)

Attribution

Modified by Melissa Ha from Community Ecology from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 3.3.1.1.2: Herbivory is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **6.1.1.2: Herbivory** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

3.3.1.1.3: Parasitism

Parasitism occurs when one organism (the **parasite**) takes nutrients from another (the **host**). The host is usually weakened by the parasite as it siphons resources the host would normally use to maintain itself. Parasites do not necessarily kill their hosts. When they do, it is often a slow process, allowing the parasite time to complete its reproductive cycle before it or its offspring are able to spread to another host. A parasite may remain attached to the same host for its full lifespan, but some parasites have complex life cycles involving multiple host species. For example, a tapeworm causes disease in humans when contaminated and under-cooked meat such as pork, fish, or beef is consumed. Figure 3.3.1.1.3.*a* illustrates the life cycle of the pork tapeworm. The tapeworm can live inside the intestine of the host for several years, benefiting from the host's food, and it may grow to be over 50 feet long by adding segments. The parasite moves from one host species to a second host species in order to complete its life cycle.

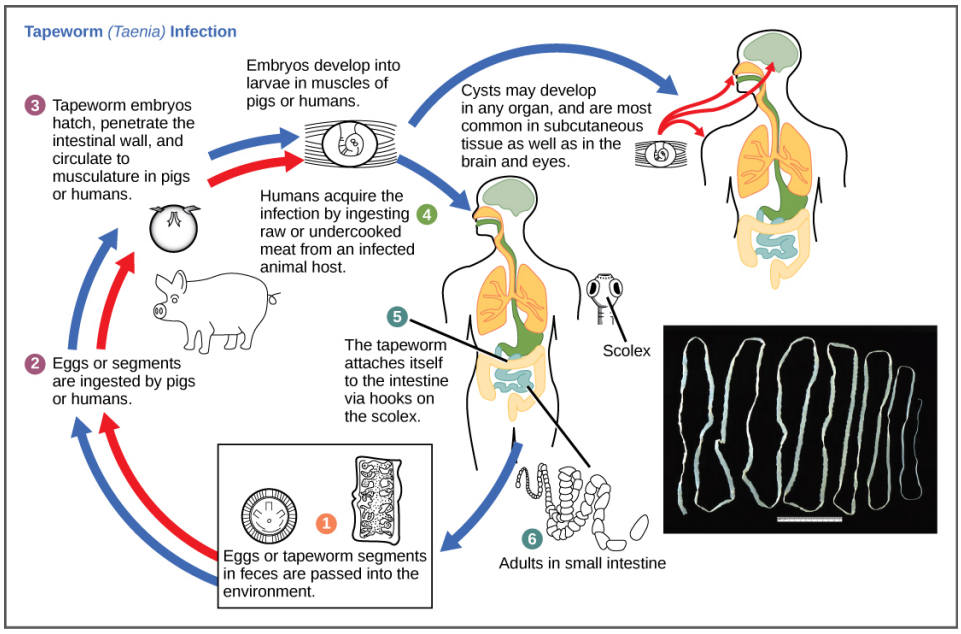


Figure 3.3.1.1.3.*a*: This diagram shows the life cycle of the tapeworm, a human worm parasite. In the first step, eggs or tapeworm segments from human feces contaminate the environment. Second, pigs ingest these eggs or segments. Third, the eggs mature into larvae, which encyst in the muscle tissue of the pig. It is also possible for humans to develop cysts if they ingest eggs. Fourth, humans become infected when they eat pork that is undercooked. Fifth, the larvae mature into adults, attaching themselves to the intestine wall. Sixth, adult tapeworms can grow many segments and become quite long. They produce the eggs or segments that continue the life cycle. (credit: modification of work by CDC)

Parasites infect many types of organisms, including other animals and plants. For example, fleas and roundworms are common dog parasites. Plants can be infected by fungi, bacteria, and viruses; there are also plants that parasitize other plants (figure 3.3.1.1.3.*b*). Even bacteria can be parasitized by viruses called bacteriophages.



Figure 3.3.1.1.3.*b*: Instead of conducting its own photosynthesis, the parasitic plant dodder siphons nutrients from its host plant. Image by Scot Nelson (public domain).

Attribution

Modified by Melissa Ha from Community Ecology from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 3.3.1.1.3: Parasitism is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **6.1.1.3: Parasitism** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.3.1.1.4: Food Chains and Food Webs

Trophic interactions in a community can be represented by diagrams called **food chains** and **food webs**. Before discussing these representations in detail, we must first review the basics of energy. Energy flows through a community as a result of trophic interactions.

Energy

Virtually every task performed by living organisms requires energy. In general, **energy** is defined as the ability to do work, or to create some kind of change. Energy exists in different forms. Examples include light energy, kinetic energy, heat energy, potential energy, and chemical energy.

When an object is in motion, there is energy associated with that object. Think of a wrecking ball. Even a slow-moving wrecking ball can do a great deal of damage to other objects. Energy associated with objects in motion is called **kinetic energy**. **Heat energy** is the energy of motion in **matter** (anything that takes up space and has mass) and is considered a type of kinetic energy. The warmer the substance, the faster its molecules are moving. The rapid movement of molecules in the air, a speeding bullet, and a walking person all have kinetic energy. Now what if that same motionless wrecking ball is lifted two stories above ground with a crane? If the suspended wrecking ball is not moving, is there energy associated with it? The answer is yes. The energy that was required to lift the wrecking ball did not disappear, but is now stored in the wrecking ball by virtue of its position and the force of gravity acting on it. This type of energy is called **potential energy**. If the ball were to fall, the potential energy would be transformed into kinetic energy until all of the potential energy was exhausted when the ball rested on the ground. Wrecking balls also swing like a pendulum; through the swing, there is a constant change of potential energy (highest at the top of the swing) to kinetic energy (highest at the bottom of the swing). Other examples of potential energy include the energy of water held behind a dam or a person about to skydive out of an airplane (figure 3.3.1.1.4.*a*).



Figure 3.3.1.1.4.*a*: Still water has potential energy; moving water, such as in a waterfall or a rapidly flowing river, has kinetic energy. (credit “dam”: modification of work by “Pascal”/Flickr; credit “waterfall”: modification of work by Frank Gualtieri)

Potential energy is not only associated with the location of matter, but also with the structure of matter. **Chemical energy** is an example of potential energy that is stored in molecules. When molecules that are higher energy and less stable react to form products that are lower energy and more stable, this stored energy is released. Chemical energy is responsible for providing living cells with energy from food.

To appreciate the way energy flows into and out of biological systems, it is important to understand two of the physical laws that govern energy. The **first law of thermodynamics** states that the total amount of energy in the universe is constant and conserved. In other words, there has always been, and always will be, exactly the same amount of energy in the universe. Energy exists in many different forms. According to the first law of thermodynamics, energy may be transferred from place to place or transformed into different forms, but it cannot be created or destroyed. The transfers and transformations of energy take place around us all the time. Light bulbs transform electrical energy into light and heat energy. Gas stoves transform chemical energy from natural gas into heat energy. Plants perform one of the most biologically useful energy transformations on earth: that of converting the energy of sunlight to chemical energy stored within biological molecules, such as sugars (figure 3.3.1.1.4.*b*).

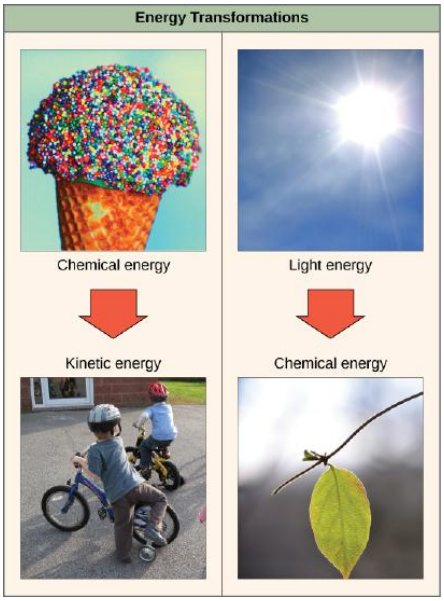


Figure 3.3.1.1.4.*b*: Shown are some examples of energy transferred and transformed from one system to another and from one form to another. The food we consume, represented by the ice cream cone, provides our cells with the chemical energy required to carry out bodily functions. This can be converted into kinetic energy (the energy of motion), which would be needed to ride a bike. Leaves conduct photosynthesis, converting light energy from the sun to chemical energy. (credit "ice cream": modification of work by D. Sharon Pruitt; credit "kids": modification of work by Max from Providence; credit "leaf": modification of work by Cory Zanker)

The challenge for all living organisms is to obtain energy from their surroundings in forms that are usable to perform cellular work. A living cell's primary tasks of obtaining, transforming, and using energy to do work may seem simple. However, the **second law of thermodynamics** explains why these tasks are harder than they appear. All energy transfers and transformations are never completely efficient. In every energy transfer, some amount of energy is lost in a form that is unusable. In most cases, this form is heat energy. For example, when a light bulb is turned on, some of the energy being converted from electrical energy into light energy is lost as heat energy. Likewise, some energy is lost as heat energy during the metabolic reactions that occur in organisms.

The concept of order and disorder relates to the second law of thermodynamics. The more energy that is lost by a system to its surroundings, the less ordered and more random the system is. Scientists refer to the measure of randomness or disorder within a system as **entropy**. High entropy means high disorder and low energy. Living things are highly ordered, requiring constant energy input to be maintained in a state of low entropy.

Energy Flow

Cells run on the chemical energy found mainly in carbohydrate molecules, and the majority of these molecules are produced by one process: **photosynthesis**. Through photosynthesis, certain organisms convert solar energy (sunlight) into chemical energy, which is then used to build carbohydrate molecules (figure 3.3.1.1.4.*c*). The energy that is harnessed from photosynthesis enters the communities continuously and is transferred from one organism to another. Therefore, directly or indirectly, the process of photosynthesis provides most of the energy required by living things on Earth. See the Carbon Cycle and Photosynthesis in OpenStax *Concepts of Biology* for more details about photosynthesis.

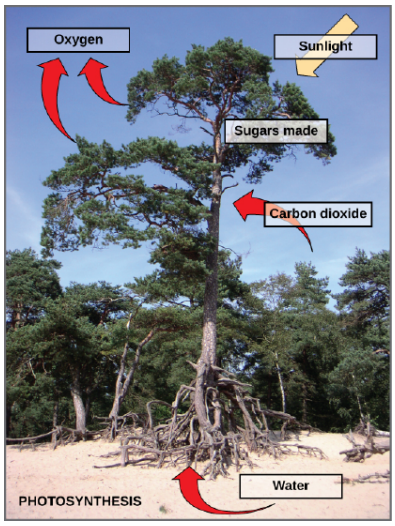


Figure 3.3.1.1.4.*c*: Photosynthesis uses solar energy, carbon dioxide, and water to release oxygen and to produce energy-storing sugar molecules. Image and caption by OpenStax (CC-BY). Access for free at openstax.org.

Organisms that conduct photosynthesis (such as plants, algae, and some bacteria), and organisms that synthesize sugars through other means are called **producers**. Without these organisms, energy would not be available to other living organisms, and life would not be possible. **Consumers**, like animals, fungi, and various microorganisms depend on producers, either directly or indirectly. For example, a deer obtains energy by eating plants. A wolf eating a deer obtains energy that originally came from the plants eaten by that deer (figure 3.3.1.1.4.*d*). Using this reasoning, all food eaten by humans can be traced back to producers that carry out photosynthesis (figure 3.3.1.1.4.*e*).



Figure 3.3.1.1.4.*d*: The energy stored in carbohydrate molecules from photosynthesis passes through the food chain. The predator that eats these deer is getting energy that originated in the photosynthetic vegetation that the deer consumed. (credit: Steve VanRiper, U.S. Fish and Wildlife Service)

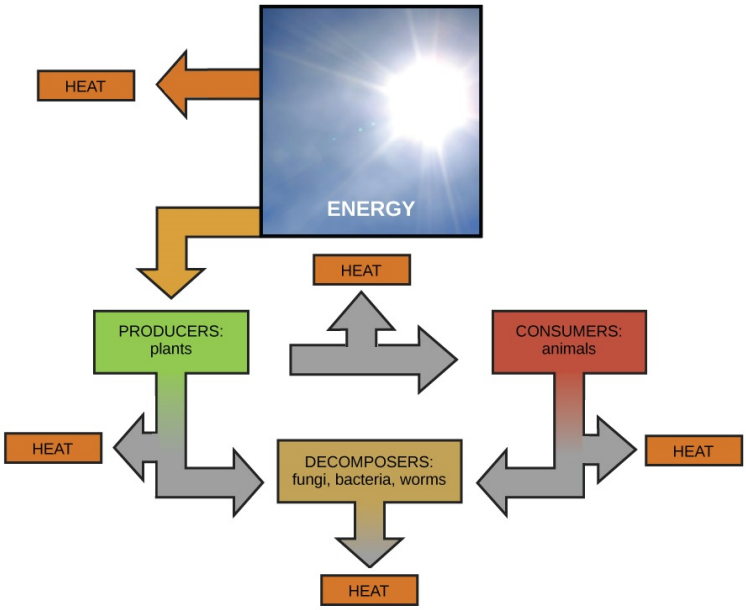


Figure 3.3.1.1.4.*e*: Ultimately, most life forms get their energy from the sun. This flow chart shows energy from the sun being captured by producers, such as plants, through photosynthesis. The energy is transferred to the consumers of the producers, such as animals. Energy can be obtained from producers directly (herbivores eat plants) or indirectly (carnivores eat herbivores). Decomposers eventually breakdown of dead organisms, including plant and animal material, and contribute to the nutrient pool. Fungi and bacteria are decomposers, and worms are detritivores (not shown). During each energy transfer, some of the energy in the system is lost as heat.

Consumers can be classified based on whether they eat animal or plant material (figure 3.3.1.1.4.*f*). Consumers that feed exclusively on animals are called **carnivores**. Lions, tigers, snakes, sharks, sea stars, spiders, and ladybugs are all carnivores. **Herbivores** are consumers that feed exclusively on plant material, and examples include deer, koalas, some bird species, crickets, and caterpillars. Herbivores can be further classified into **frugivores** (fruit-eaters), **granivores** (seed eaters), **nectivores** (nectar feeders), and **folivores** (leaf eaters). Consumers that eat both plant and animal material are considered **omnivores**. Humans, bears, chickens, cockroaches, and crayfish are examples of omnivores.



Figure 3.3.1.1.4.*f*: Carnivores like the lion (top left) eat primarily meat. The ladybug (lower left) is also a carnivore that consumes small insects called aphids. Herbivores, like the mule deer (middle) eat primarily plant material. Omnivores like the bear (top right) and crayfish (bottom right) eat both plant and animal based food. Lion by Kevin Pluck; ladybug by Jon Sullivan; mule deer by Bill Ebbesen; bear by Dave Menke; crayfish by Jon Sullivan. All from OpenStax (CC-BY). Access for free at openstax.org.

Dead producers and consumers are eaten by **detritivores** (which ingest dead tissues) and **decomposers** (which further break down these tissues into simple molecules by secreting digestive enzymes). Invertebrate animals, such as worms and millipedes, are examples of detritivores, whereas fungi and certain bacteria are examples decomposers.

Food Chains

A **food chain** is a linear sequence of organisms through which nutrients and energy pass as one organism eats another. Each organism in a food chain occupies a specific **trophic level** (energy level), its position in the food chain. The first trophic level in the food chain is the producers. The **primary consumers** (the herbivores that eat producers) are the second trophic level. Next are higher-level consumers. Higher-level consumers include **secondary consumers** (third trophic level), which are usually carnivores that eat the primary consumers, and **tertiary consumers** (fourth trophic level), which are carnivores that eat other carnivores. In the Lake Ontario food chain, shown in figure 3.3.1.1.4.*g*, the Chinook salmon is the apex consumer at the top of this food chain. Some communities have additional trophic levels (quaternary consumers, fifth order consumers, etc.). Finally, detritivores and decomposers break down dead and decaying organisms from any trophic level. There is a single path through a food chain.

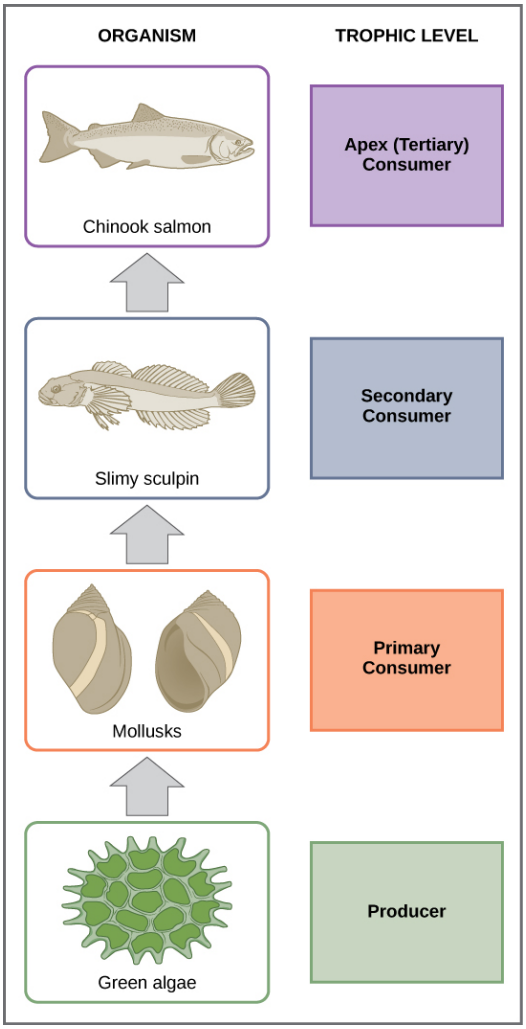


Figure 3.3.1.1.4.*g*: These are the trophic levels of a food chain in Lake Ontario at the United States–Canada border. Energy and nutrients flow from photosynthetic green algae (producers) at the base to the primary consumers, which are mollusks, or snails. The secondary consumers are small fish called slimy sculpin. The tertiary and apex consumer is Chinook salmon. Detritivores and decomposers are not shown. (credit: modification of work by National Oceanic and Atmospheric Administration/NOAA)

One major factor that limits the number of steps in a food chain is energy. Only about 10% of the energy in one trophic level is transferred to the next trophic level. This is because much energy is lost as heat during transfers between trophic levels or to decomposers due to the second law of thermodynamics. Thus, after four to six trophic energy transfers, the amount of energy remaining in the food chain may not be great enough to support viable populations at higher trophic levels (also see Community Productivity and Transfer Efficiency).

Certain environmental toxins can become more concentrated as they move up the food chain, with the highest concentrations occurring in the top consumers, a process called **biomagnification**. Essentially, a top consumer ingests all the toxins that had previously accumulated in the bodies of the organisms at the lower trophic levels. This explains why frequently eating certain fish, like tuna or swordfish, increases your exposure to mercury, a toxic heavy metal.

Food Webs

While food chains are simple and easy to analyze, there is a one problem when using food chains to describe most communities. Even when all organisms are grouped into appropriate trophic levels, some of these organisms can feed at more than one trophic level. In addition, species feed on and are eaten by more than one species. In other words, the linear model of trophic interactions, the food chain, is a hypothetical and overly simplistic representation of community structure. A holistic model—which includes all the interactions between different species and their complex interconnected relationships with each other and with the environment

—is a more accurate and descriptive model. A **food web** is a concept that accounts for the multiple trophic interactions between each species (figure 3.3.1.1.4.*h* and i).

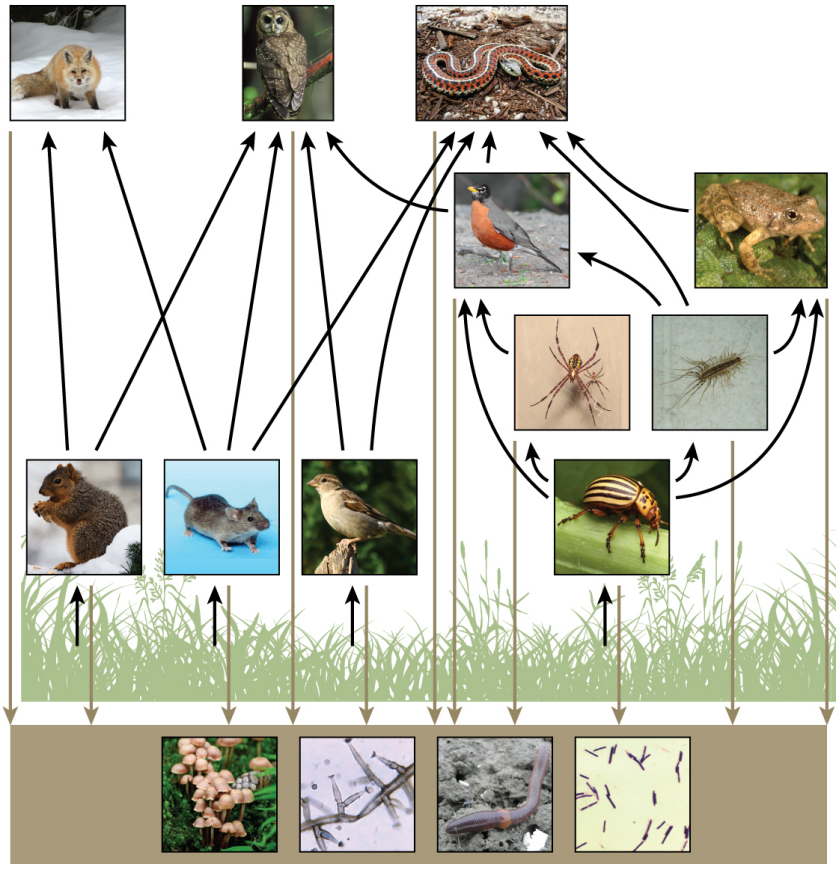


Figure 3.3.1.1.4.*h*: This food web shows the interactions among organisms across trophic levels. Arrows point from an organism that is consumed to the organism that consumes it and represent energy transfer. The producers (plants) capture energy from the sun. The level above the producers shows the primary consumers that eat the producers. Some examples are squirrels, mice, seed-eating birds, and beetles. Spiders and centipedes eat beetles. Robins eat beetles, spiders, and centipedes, and toads eat beetles and centipedes. Foxes eat squirrels and mice; owls eat squirrels, mice, seed-eating birds, and robins; and snakes eat mice, seed-eating birds, robins, centipedes, and toads. Some consumers are in between trophic levels because they eat a combination of primary, secondary, and/or tertiary consumers. All the producers and consumers eventually become nourishment for the decomposers (mushrooms, mold, and bacteria) and detritivores (earthworms) in the soil, which are depicted at the bottom of the illustration. (credit “fox”: modification of work by Kevin Bacher, NPS; credit “owl”: modification of work by John and Karen Hollingsworth, USFWS; credit “snake”: modification of work by Steve Jurvetson; credit “robin”: modification of work by Alan Vernon; credit “frog”: modification of work by Alessandro Catenazzi; credit “spider”: modification of work by “Sanba38″/Wikimedia Commons; credit “centipede”: modification of work by “Bauerph”/Wikimedia Commons; credit “squirrel”: modification of work by Dawn Huczek; credit “mouse”: modification of work by NIGMS, NIH; credit “sparrow”: modification of work by David Friel; credit “beetle”: modification of work by Scott Bauer, USDA Agricultural Research Service; credit “mushrooms”: modification of work by Chris Wee; credit “mold”: modification of work by Dr. Lucille Georg, CDC; credit “earthworm”: modification of work by Rob Hille; credit “bacteria”: modification of work by Don Stalons, CDC)

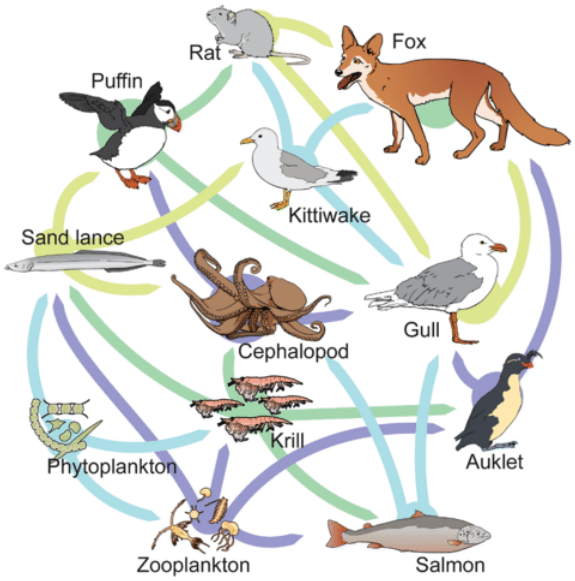


Figure 3.3.1.1.4.*i*: In this food web, each organism may have multiple food sources or be eaten by multiple species. Phytoplankton are the primary producers. They are consumed by sand lances, krill, and zooplankton. Sand lances are consumed by puffins, kittiwake, and cephalopods. Krill are consumed by the sand lance, cephalopods, auklets, and salmon. Zooplankton are consumed by the sand lance, krill, auklets, and salmon. Puffins are eaten by rats and gulls. Kittiwakes are eaten by rats, foxes, and gulls. Cephalopods are consumed by puffins and gulls. Auklets are consumed by gulls and foxes. Salmon are consumed by cephalopods and gulls. Rats are consumed by foxes and gulls. Gulls are consumed by foxes. Detritivores and decomposers are not shown. Image by Mariana Ruiz Villarreal (LadyofHats) for CK-12 Foundation (CC-BY-NC).

The trophic level of each species in a food web is not necessarily a whole number. In figure 3.3.1.1.4.*i*, phytoplankton are the primary producers (trophic level 1). Zooplankton only feed on phytoplankton, making them primary consumers (trophic level 2). Determining the trophic level of the other species is more complex. For example, krill eat both phytoplankton and zooplankton. If krill only ate phytoplankton they would primary consumers (trophic level 2). If they ate only zooplankton, they would be secondary consumers (trophic level 3). Since, krill consume both, their trophic level is 2.5.

Community Productivity and Transfer Efficiency

The rate at which photosynthetic producers incorporate energy from the sun is called **gross primary productivity**. In a cattail marsh, plants only trap 2.2% of the energy from the sun that reaches them. Three percent of the energy is reflected, and another 94.8% is used to heat and evaporate water within and surrounding the plant. However, not all of the energy incorporated by producers is available to the other organisms in the food web because producers must also grow and reproduce, which consumes energy. At least half of the 2.2% trapped by cattail marsh plants is used to meet the plants own energy needs.

Net primary productivity is the energy that remains in the producers after accounting for the metabolic needs of the producers and heat loss. The net productivity is then available to the primary consumers at the next trophic level. One way to measure net primary productivity is to collect and weigh the plant material produced on a m^2 (about 10.7 ft^2) of land over a given interval. One gram of plant material (e.g., stems and leaves), which is largely the carbohydrate cellulose, yields about 4.25 kcal of energy when burned. Net primary productivity can range from 500 kcal/m^2/yr in the desert to 15,000 kcal/m^2/yr in a tropical rain forest.

In an aquatic community in Silver Springs, Florida, the gross primary productivity (total energy accumulated by the primary producers) was 20,810 kcal/m^2/yr (figure 3.3.1.1.4.*j*). The net primary productivity (energy available to consumers) was only 7,632 kcal/m^2/yr after accounting for energy lost as heat and energy require to meet the producer's metabolic needs.

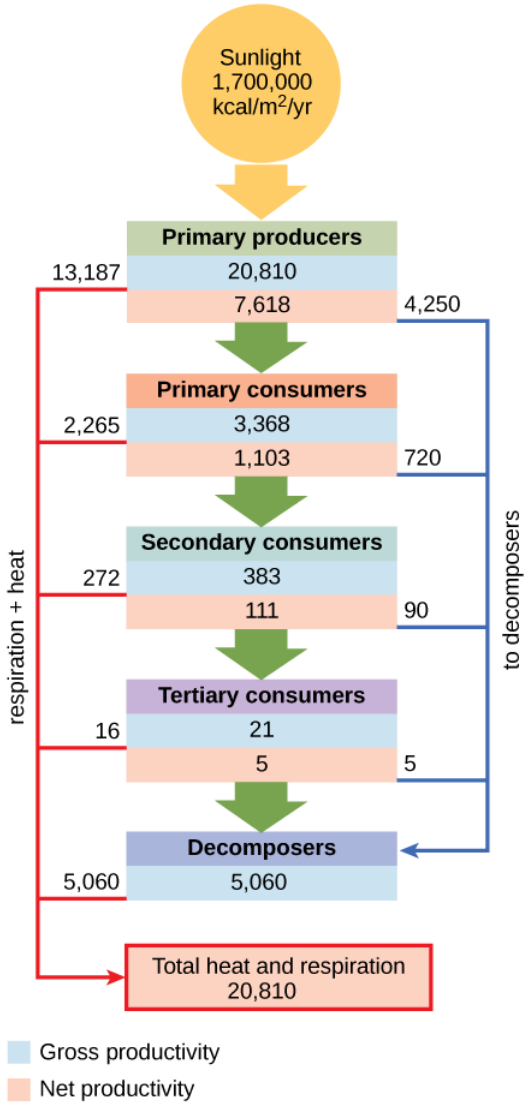


Figure 3.3.1.1.4.*j*: The flow of energy through a spring ecosystem in Silver Springs, Florida. The gross productivity (blue) for each trophic level is listed just above the net productivity (red). Notice that the energy decreases with each increase in trophic level. The energy content of primary producers (gross productivity) is 20,810 kcal/m^2/yr. The gross productivity of primary consumers is much smaller, about 3,368 kcal/m^2/yr. The gross productivity of secondary consumers is 383 kcal/m^2/yr, and the gross productivity of tertiary consumers is only 21 kcal/m^2/yr. The net productivity of each trophic level is less than the gross productivity because some energy is used to meet metabolic needs (respiration), and some energy is lost as heat. For example, the net productivity of primary consumers was 1,103 kcal/m^2/yr, only about a third of the gross productivity. Some of the energy in each trophic level (a total of 5,060 kcal/m^2/yr) is transferred to decomposers. All of the energy initially captured by primary producers (20,810 kcal/m^2/yr) is eventually released from the system.

Only a fraction of the energy captured by one trophic level is assimilated into biomass, which makes it available to the next trophic level. Assimilation is the biomass of the present trophic level after accounting for the energy lost due to incomplete ingestion of food, energy used to conduct work by that trophic level, and energy lost as waste. Incomplete ingestion refers to the fact that some consumers eat only a part of their food. For example, when a lion kills an antelope, it will eat everything except the hide and bones. The lion is missing the energy-rich bone marrow inside the bone, so the lion does not make use of all the calories its prey could provide. In Silver Springs, only 1103 kcal/m^2/yr from the 7618 kcal/m^2/yr of energy available to primary consumers was assimilated into their biomass. (The **trophic level transfer efficiency** between the first two trophic levels was approximately 14.8 percent.)

An animal's source of heat influences its energy needs. **Ectotherms**, such as invertebrates, fish, amphibians, and reptiles, rely on external sources for body heat, and **endotherms**, such as birds and mammals, rely on internally generated heat. Generally,

ectotherms require less of the energy to meet their metabolic needs and than endotherms do, and therefore, many endotherms have to eat more often than ectotherms.

The inefficiency of energy use by endotherms has broad implications for the world's food supply. It is widely accepted that the meat industry uses large amounts of crops to feed livestock, and because a low percentage of this is assimilated into biomass, much of the energy from animal feed is lost. For example, it costs about 1¢ to produce 1000 dietary calories (kcal) of corn or soybeans, but approximately $0.19 to produce a similar number of calories growing cattle for beef consumption. The same energy content of milk from cattle is also costly, at approximately $0.16 per 1000 kcal. Thus, there has been a growing movement worldwide to promote the consumption of non-meat and non-dairy foods so that less energy is wasted feeding animals for the meat industry.

Attribution

Modified by Melissa Ha from the following sources:

- Energy, Energy Enters Ecosystems Through Photosynthesis, and Energy Flow Through Ecosystems from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Ecology of Ecosystems, Energy Flow through Ecosystems, and Digestive Systems from *General Biology/Biology 2e* by OpenStax (licensed under CC-BY). Access for free at openstax.org.
- Ecosystem Productivity from *Biology* by John W. Kimball (licensed under CC-BY)

This page titled 3.3.1.1.4: Food Chains and Food Webs is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **6.1.1.4: Food Chains and Food Webs** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.3.1.2: Competition

Resources are often limited within a habitat and multiple species may compete to obtain them. **Competition** occurs when organisms use the same resources, and one or both organisms is harmed. Competing species are often occupy the same trophic level and compete for food, but species on different trophic levels could still compete for space, water, etc. **Intraspecific competition** occurs within a species. For example, most penguins defend territories from other individuals of the same species because they compete for suitable habitat and other resources (figure 3.3.1.2.*a*). **Interspecific competition** occurs between different species. For example, the **invasive** vine, kudzu, competes with trees in the southeastern United States for light (figure 3.3.1.2.*b*). Kudzu is considered invasive in this region because it occurs outside of its historical range (Asia and Australia) and causes ecological harm.



Figure 3.3.1.2.*a*: Adélie penguins at Cape Adare in Ross Sea, Antarctica. These penguins defend small territories in which they nest. Image by Brocken Inaglory (CC-BY-SA).



Figure 3.3.1.2.*b*: Kudzu is a vine that covers trees, shrubs, and other structures. It can fully block its competitors' access to light, eventually killing them. Image by Katie Ashdown (CC-BY).

Ecologists have come to understand that all species have an **ecological niche**: the unique set of resources used by a species, which includes its interactions with other species. The **competitive exclusion principle** states that two species cannot occupy the exact same niche in a habitat. In other words, different species cannot coexist in a community if they are competing for all the same resources. Competition harms one or both competitors because it wastes energy. In an experimental test of competitive exclusion, two species of the freshwater microbe *Paramecium* were cultured separately and together. When cultured separately, both species

reproduced, and the number of cells (individuals) increased. However, when the two species were grown together, one species (*P. aurelia*) grew, and the the other species (*P. caudatum*) was eliminated. Because both species occupied the same ecological niche, they could not coexist. *Paramecium aurelia* was the superior competitor in this case (figure 3.3.1.2.*c*). Competitive exclusion may be avoided if a population evolves to make use of a different resource, a different area of the habitat, or feeds during a different time of day. This is called **resource partitioning**. The two species are then said to occupy different microniches. These species coexist by minimizing direct competition.

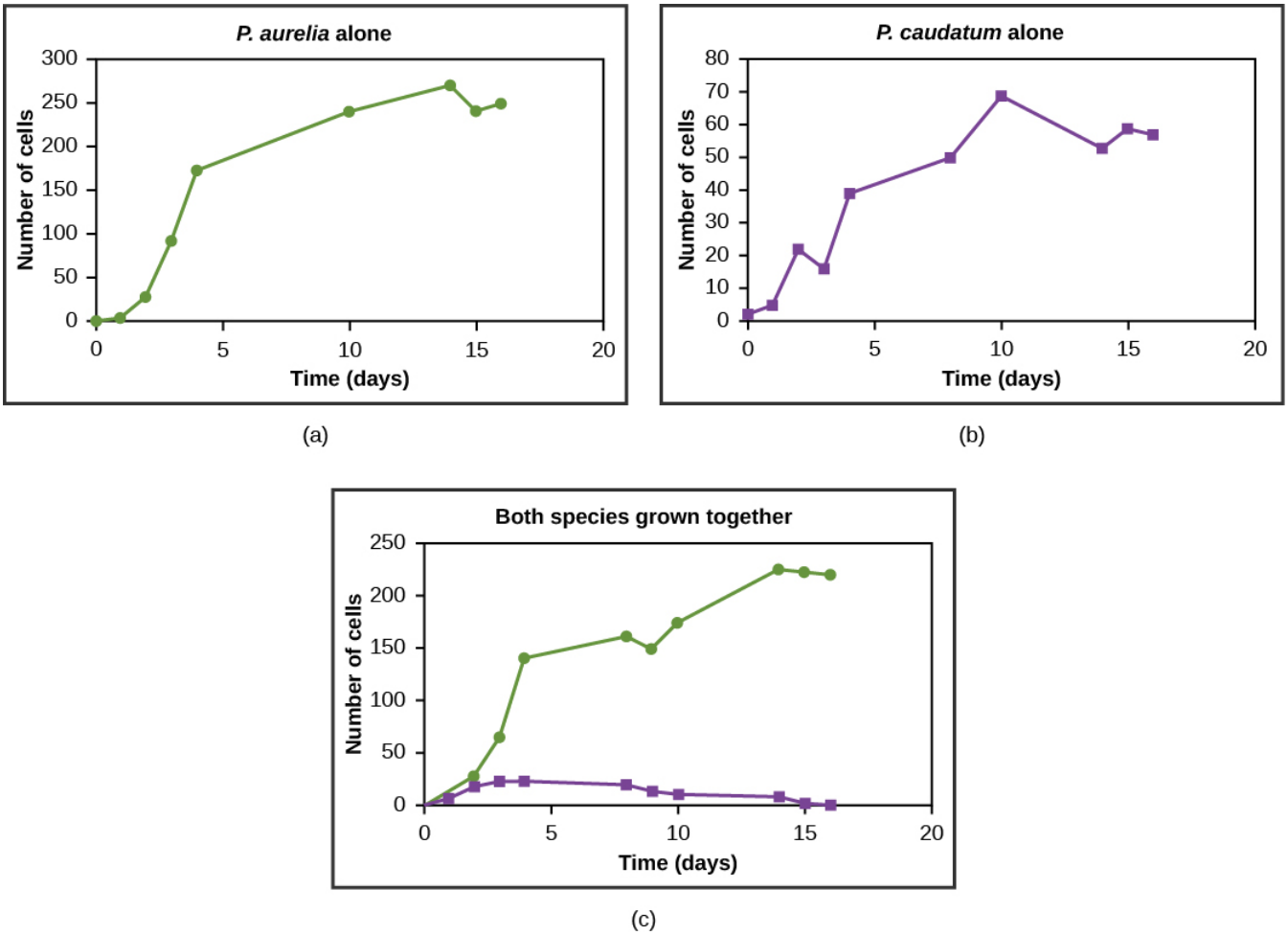


Figure 3.3.1.2.*c*: Graphs a, b, and c all plot number of cells versus time in days. When *Paramecium aurelia* and *P. caudatum* were grown individually in the laboratory (a and b), the two species both reproduced and reached a relatively high cell density. When the two species were placed together in the same test tube (habitat, c), *P. aurelia* (top, green) reached nearly the same cell density as it exhibited when grown alone, but *P. caudatum* (bottom, purple) hardly grew at all, and eventually its population drops to zero. *Paramecium aurelia* outcompeted *P. caudatum* for food, leading to the latter's eventual extinction.

Attribution

Modified by Melissa Ha from Community Ecology from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 3.3.1.2: Competition is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **6.1.2: Competition** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

3.3.1.3: Facilitation

Interactions in which one or both species benefit and neither is harmed are called facilitation. There are two types of facilitation: commensalism and mutualism.

Commensalism

Commensalism is a type of facilitation that occurs when one species benefits from an interaction, while the other neither benefits or is harmed. Many potential commensal relationships are difficult to identify because it is difficult to demonstrate that one partner is unaffected by the presence of the other. Birds nesting in trees provide an example of a commensal relationship (figure 3.3.1.3.a). The tree is not harmed by the presence of the nest among its branches. The nests are light and produce little strain on the structural integrity of the branch, and most of the leaves, which the tree uses to get energy by photosynthesis, are above the nest so they are unaffected. The bird, on the other hand, benefits greatly. If the bird had to nest in the open, its eggs and young would be vulnerable to predators.



Figure 3.3.1.3.a: The Southern Masked-Weaver is starting to make a nest in a tree in Zambezi Valley, Zambia. This is an example of a commensal relationship, in which one species (the bird) benefits, while the other (the tree) neither benefits nor is harmed. (credit: "Hanay"/Wikimedia Commons)

Another example or a commensal relationship involves the Little Blue Heron and the White Ibis, which are both wading birds. The Little Blue Heron catches more fish in the presence of the White Ibis, but the White Ibis is unaffected. Interestingly, Little Blue Herons attempt to catch fish more often in the presence of the species, but the success rate of their attempts does not change. Nevertheless, more frequent attempts still increases the total number of fish caught. The White Ibis may make fish more visible to Little Blue Herons, causing changes in their behavior (figure 3.3.1.3.b).



Figure 3.3.1.3.b: The Little Blue Heron (left) and the White Ibis (right) have a commensal relationship. Image by Russ (CC-BY).

Mutualism

In a **mutualism**, both species benefit from their interaction. For example, pollinators, such as bees, butterflies, and hummingbirds, benefit because they eat the collect pollen and/or nectar that they collect from flowers. The plants also benefit because their pollen is dispersed to other plants, allowing them to reproduce. Both the pollinators and the plants benefit, demonstrating a mutualism (figure 3.3.1.3.c). Clownfish and anemones are another example of mutualisms. Clownfish gain protection from living among anemones. In return, clownfish clean the anemones and scare away predators. In addition, their waste provides nutrients (figure 3.3.1.3.d).



Figure 3.3.1.3.c: A male Broad-Tailed Hummingbird visits a scarlet gilia flower at the Rocky Mountain Biological Laboratory. The hummingbird gains food (nectar) while aiding the gilia flower with reproduction. Image by David W. Inouye (CC-BY).



Figure 3.3.1.3.d: These clownfish have a mutual relationship with the sea anemones in which they live. Image by Samuel Chow (CC-BY).

Attribution

Modified by Melissa Ha from Community Ecology from *Environmental Biology* by Matthew R. Fisher's (licensed under CC-BY)

This page titled 3.3.1.3: Facilitation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **6.1.3: Facilitation** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.3.1.4: Symbiosis

Some biological interactions are brief, such as predation. In others, species are closely associated for long periods. Such associations are called **symbiotic** ("living together"). One species always benefits in a symbiosis, but the other may be harmed (parasitism), unaffected (commensalism), or benefited (mutualism). Some experts considered all parasitisms, commmensalisms, and mutualisms to be symbiotic, but others only consider interactions in which species are living together in a close association (such as when one species lives on or in the other) as symbiotic.

Termites have a symbiotic relationship with microorganisms called protozoa that live in the insect's gut (figure 3.3.1.4.*a*-a). The termite benefits from the ability of the protozoa to digest cellulose, a carbohydrate important in plant structure. However, the protozoa are able to digest cellulose only because of the presence of symbiotic bacteria within their cells that produce the cellulase enzyme. The termite itself cannot do this; without the protozoa, it would not be able to obtain energy from its food (cellulose from the wood it chews and eats). The protozoa benefit by having a protective environment and a constant supply of food from the wood chewing actions of the termite. In turn, the protists benefit from the enzymes provided by their bacterial endosymbionts, while the bacteria benefit from a doubly protective environment and a constant source of nutrients from two hosts.

Lichen are a symbiotic relationship between a fungus and algae (or a photosynthetic bacteria; Figure 3.3.1.4.*a*-b). In lichen, the algal cells are fully surrounded by the fungus. The sugars produced by the algae through photosynthesis provide nourishment for both organisms. The physical structure of the fungus protects the algae from the elements and makes certain nutrients in the atmosphere more available to the algae. The algae of lichens can sometimes live independently given the right environment, but many of the fungal partners are unable to live on their own. Note that while both lichen and the example involving termites mutually benefit the species involved, symbioses are not always mutualisms.



Figure 3.3.1.4.*a*: (a) Termites form a symbiotic relationship with protozoa living in their guts. Their association which allows both organisms to obtain energy from the cellulose the termite consumes. (b) Lichen is a fungus that has symbiotic photosynthetic algae living in close association. (credit a: modification of work by Scott Bauer, USDA; credit b: modification of work by Cory Zanker)

The remora and the shark is a commensal symbiosis. The top fin of the remora is modified into a sucker with which it forms an attachment to the shark. When the shark feeds, the remora picks up scraps. Additionally, the remora gains transportation and protection. The shark makes no attempt to prey on the remora and is unaffected by it (figure 3.3.1.4.*b*).

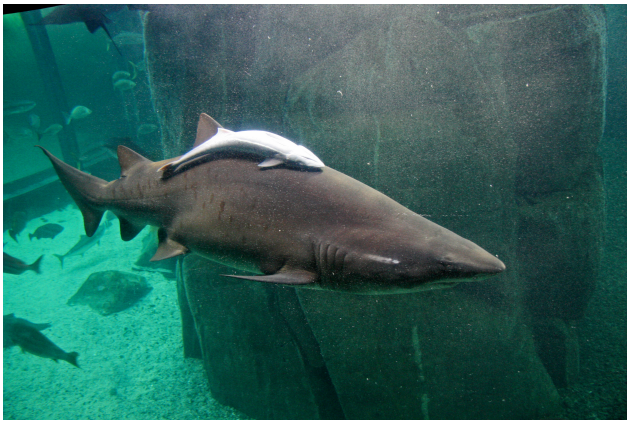


Figure 3.3.1.4.*b*: A remora attached to the top of a shark in an aquarium. This is an example of a symbiosis because the remora lives in close association with the shark. Image by Brian Snelson (CC-BY)

There are many additional examples of symbiosis. Coral are small animals that harbor photosynthetic microorganisms called dinoflagellates. (The corals group together and deposit hardened calcium carbonate skeletons.) **Mycorrhizae** are fungi that surround or live within plant root cells and help plants absorb nutrients from the soil in exchange for sugars. In fact truffles are often found in oak forests because the fungus that produce this spore-forming body establishes mycorrhizae on oak roots. **Root nodules** are structures on the roots of legume plants that house nitrogen-fixing bacteria (see The Nitrogen Cycle). **Epiphytes** are plants adapted to live on top of other plants and are common in the tropical rainforest (see Biomes). Humans have a diversity of microorganisms, called the **normal flora**, that live in and on our bodies. While we often think of bacteria as causing infection (a parasitism), the species of our normal floral exemplify commensalism and mutualism. For example, our gut microbiota aids with digestion and helps us synthesize certain vitamins.

Attributions

Modified by Melissa Ha from the following sources:

- Symbiosis from *Biology* by John W. Kimball (licensed under CC-BY)
- Community Ecology from *Environmental Biology* by Matthew R. Fisher's (licensed under CC-BY)

This page titled 3.3.1.4: Symbiosis is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **6.1.4: Symbiosis** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.3.2: Community Structure and Dynamics

Communities are complex systems that can be characterized by **community structure** (the number and size of populations and their interactions) and **community dynamics** (how the members and their interactions change over time). Understanding community structure and dynamics allows us to minimize impacts on ecosystems and manage ecological communities we benefit from.

Keystone Species

A **keystone species** is one whose presence has inordinate influence in maintaining the prevalence of various species, the ecological community's structure, and sometimes its biodiversity. *Pisaster ochraceus*, the intertidal sea star, is a keystone species in the northwestern portion of the United States (figure 3.3.2.*a*). Studies have shown that when this organism is removed from communities, mussel populations (their natural prey) increase, which completely alters the species composition and reduces biodiversity. Another keystone species is the banded tetra, a fish in tropical streams, which supplies nearly all of the phosphorus, a necessary inorganic nutrient, to the rest of the community. The banded tetra feeds largely on insects from the terrestrial ecosystem and then excretes phosphorus into the aquatic ecosystem. The relationships between populations in the community, and possibly the biodiversity, would change dramatically if these fish were to become extinct.



Figure 3.3.2.*a*: Sea stars are keystone species of the intertidal zone.

Community Dynamics

Community dynamics are the changes in community structure and composition over time, often following environmental disturbances such as volcanoes, earthquakes, storms, fires, and climate change. Communities with a relatively constant number of species are said to be at equilibrium. The equilibrium is dynamic with species identities and relationships changing over time, but maintaining relatively constant numbers. Following a disturbance, the community may or may not return to the equilibrium state.

Succession describes the sequential appearance and disappearance of species in a community over time after a severe disturbance. In **primary succession**, newly exposed or newly formed rock is colonized by living organisms. In **secondary succession**, a part of an ecosystem is disturbed and remnants of the previous community remain. In both cases, there is a sequential change in species until a more-or-less permanent community develops.

Primary Succession and Pioneer Species

Primary succession occurs when new land is formed, or when the soil and all life is removed from pre-existing land. An example of the former is the eruption of volcanoes on the Big Island of Hawaii, which results in lava that flows into the ocean and continually forms new land. From this process, approximately 32 acres of land are added to the Big Island each year. An example of pre-existing soil being removed is through the activity of glaciers. The massive weight of the glacier scours the landscape down to the bedrock as the glacier moves. This removes any original soil and leaves exposed rock once the glacier melts and retreats.

In both cases, the ecosystem starts with bare rock that is devoid of life. New soil is slowly formed as weathering and other natural forces break down the rock and lead to the establishment of hearty organisms, such as lichens and some plants, which are collectively known as **pioneer species** (figure 3.3.2.*b*) because they are the first to appear. These species help to further break down the mineral-rich rock into soil where other, less hardy but more competitive species, such as grasses, shrubs, and trees, will grow and eventually replace the pioneer species. Over time the area will reach an equilibrium state, with a set of organisms quite different from the pioneer species.



Figure 3.3.2.*b*: During primary succession in lava on Maui, Hawaii, succulent plants are the pioneer species. (credit: Forest and Kim Starr)

The relationship between pioneer species and the more competitive (late-successional) species illustrates the complexity of biotic interactions. Pioneer species facilitate the growth of the late-successional species, and this initially appears to be a commensalism. However, the late-successional species later outcompete the pioneer species, shifting the interaction to competition.

Secondary Succession

A classic example of secondary succession occurs in forests cleared by wildfire, or by clearcut logging (figure 3.3.2.*c*). Wildfires will burn most vegetation, and unless the animals can flee the area, they are killed. Their nutrients, however, are returned to the ground in the form of ash. Thus, although the community has been dramatically altered, there is a soil ecosystem present that provides a foundation for rapid recolonization.

Before the fire, the vegetation was dominated by tall trees with access to the major plant energy resource: sunlight. Their height gave them access to sunlight while also shading the ground and other low-lying species. After the fire, though, these trees are no longer dominant. Thus, the first plants to grow back are usually annual plants followed within a few years by quickly growing and spreading grasses and other pioneer species. Due, at least in part, to changes in the environment brought on by the growth of grasses and forbs, over many years, shrubs emerge along with small trees. These organisms are called intermediate species. Eventually, over 150 years or more, the forest will reach its equilibrium point and resemble the community before the fire. This equilibrium state is referred to as the **climax community**, which will remain until the next disturbance. The climax community is typically characteristic of a given climate and geology. Although the community in equilibrium looks the same once it is attained, the equilibrium is a dynamic one with constant changes in abundance and sometimes species identities.

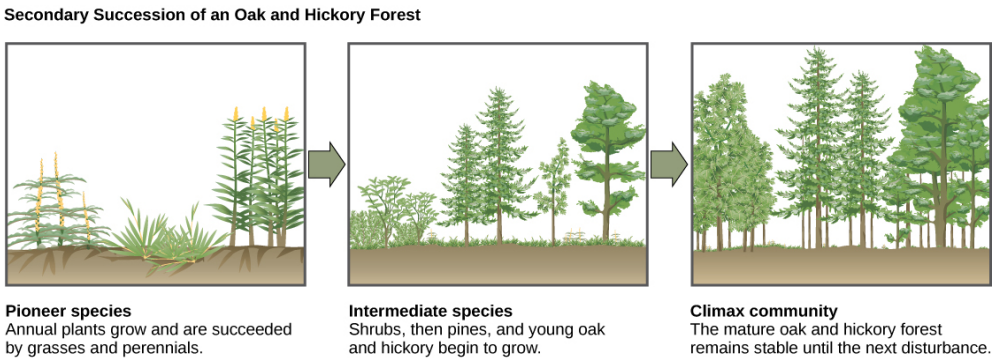


Figure 3.3.2.*c*: Succession in an oak and hickory forest. Secondary succession occurs when a disturbance (reduction in biomass) occurs, but the entire community does not die as a result of the disturbance. First, pioneer species grow. These include annual plants, which are succeeding by perennial grasses and other perennials. Next, intermediate species, grow. These include shrubs followed by pines and young oak and hickory. Finally, the climax community is established, which consists of mature oak and hickory trees. These remain intact until the next disturbance.

Attribution

Modified by Melissa Ha from Community Ecology from *Environmental Biology* by Matthew R. Fisher's (licensed under CC-BY)

This page titled 3.3.2: Community Structure and Dynamics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **6.2: Community Structure and Dynamics** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.3.3: Data Dive- Bison Impacts In Prairies

Overview

In many grassland ecosystems the activities of large grazers, like bison, play a key role in shaping grassland habitat structure and diversity. A 2017 study sought to investigate how bison in tall-grass prairies affect local plant competition. Specifically, they wanted to know if bison have a positive impact on herbaceous forbs (non-woody flower species). To investigate this, they observed plant communities in tall-grass prairies in the presence and absence of bison. The two graphs below display some of the results in this study:

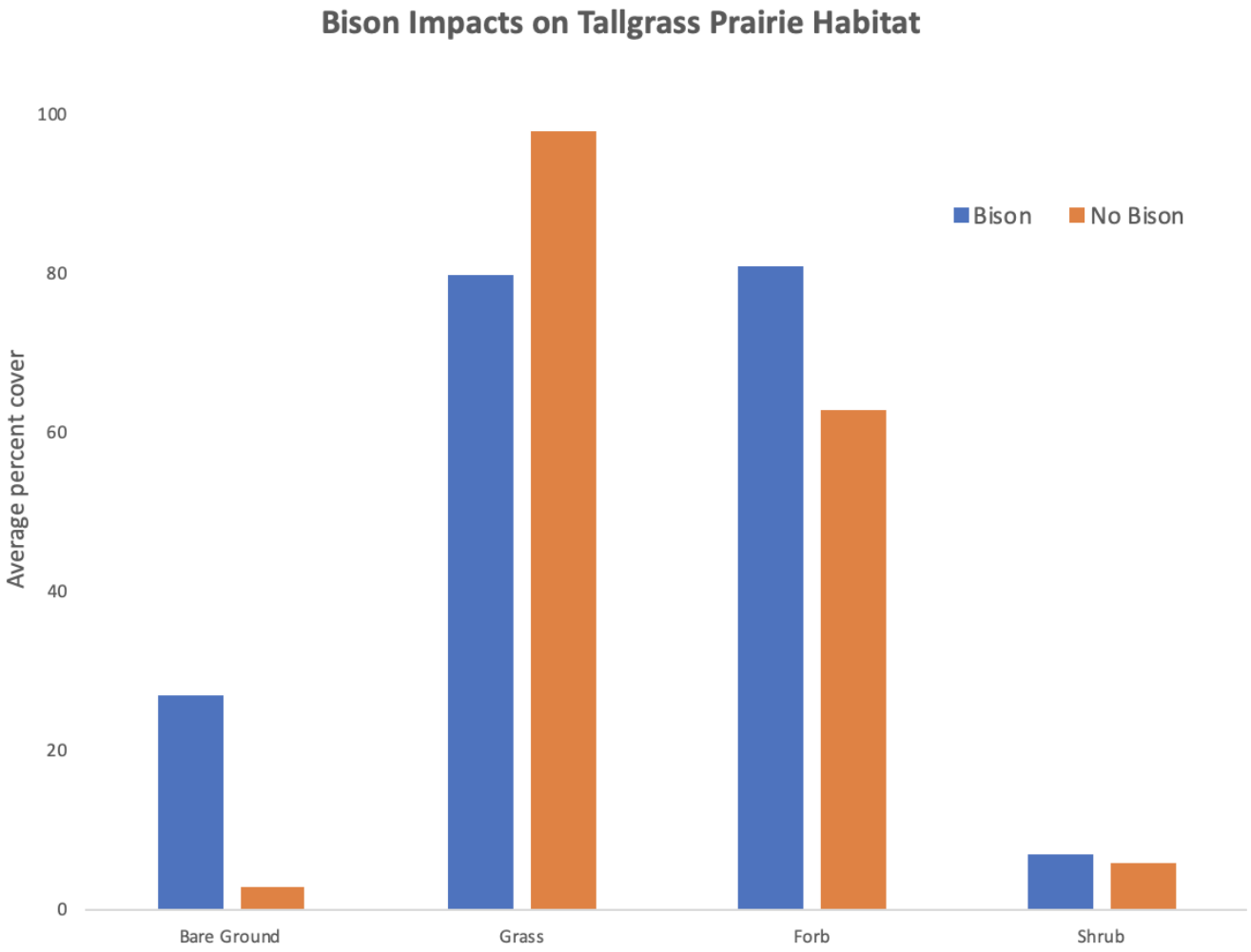


Figure 3.3.3.2a: Average percent cover (aka what percent of an area a species took up) for grasses, forbs, shrubs, and bare ground in tall-grass prairies with and without bison. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Elson A and Hartnett D, 2017.

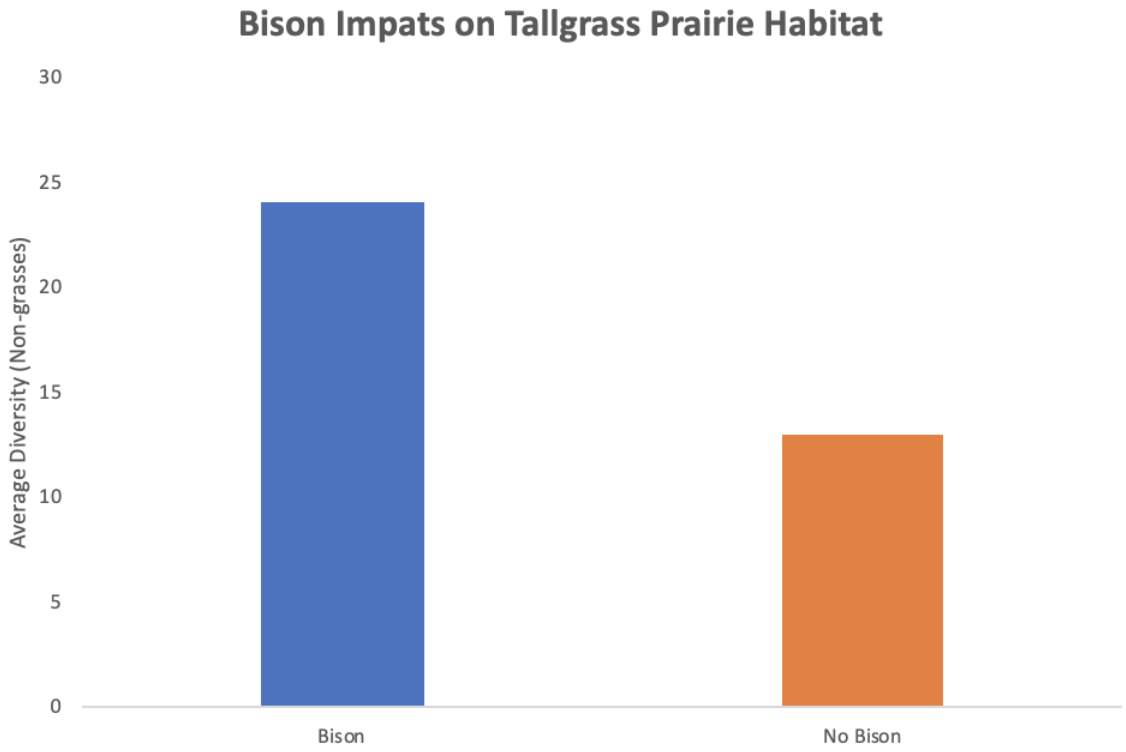


Figure 3.3.3.*b*: Average diversity of non-grasses in tall-grass prairies with and without bison. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Elson A and Hartnett D, 2017.

Questions

1. What kind of graphs are these?
2. Does the presence of bison increase diversity? How do you know this? Which graph(s) support this?
3. How does the presence of bison affect the percent cover of bare ground, grasses, forbs, and shrubs?
4. Based on the authors objectives/predictions described above, do the results of these graphs support them? Why/why not?
5. How can we use the results of this graph to inform conservation projects for tallgrass prairies in the future?
6. What does the results of this graph make you curious about?

Raw Data From Above Graph(s)

Table 3.3.3.2*a*: Raw data for average percent covers (aka what percent of an area a species took up) for grasses, forbs, shrubs, and bare ground in tall-grass prairies with and without bison. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Elson A and Hartnett D, 2017.

Status	Average Percent Cover With Bison	Average Percent Cover With No Bison
Bare Ground	27	3
Grass	80	98
Forb	81	63
Shrub	7	6

Table 3.3.3.*b*: Raw data for average diversity of non-grasses in tall-grass prairies with and without bison. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Elson A and Hartnett D, 2017.

Average Diversity (Non-Grasses) With Bison	Average Diversity (Non-Grasses) With No Bison
24	13

Attributions

Rachel Schleiger (CC-BY-NC)

This page titled 3.3.3: Data Dive- Bison Impacts In Prairies is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **6.3: Data Dive- Bison Impacts In Prairies** by Melissa Ha and Rachel Schleiger has no license indicated.

3.3.4: Review

Summary

After completing this chapter you should be able to...

- Name and distinguish among the major types of biotic interactions, providing examples of each.
- Describe how prey have adapted to reduce predation and how plants have adapted to reduce herbivory.
- Define energy and distinguish among the different types of energy.
- Explain the first and second laws of thermodynamics as they relate to trophic interactions in a community.
- Distinguish between and interpret food webs and food chains, explaining the advantages of each representation and identifying the trophic levels in each.
- Explain what limits the number of trophic levels in a community.
- Define symbiosis and provide examples.
- Explain the role of keystone species in maintaining community structure.
- Summarize the process of succession.

Communities consists of populations of different species interacting in a common area. **Biotic interactions** describe the relationships among organisms, and **trophic interactions**, including **predation**, **herbivory**, and **parasitism**, are the biotic interactions that involve one organism eating another. Trophic interactions are diagrammed through **food chains** and **food webs**.

Some biotic interactions are not trophic. For example, **competition** is a biotic interaction in which one or both organisms that use the same resources are harmed, and it can occur between individuals of the same species (**intraspecific**) or between members of different species (**interspecific**). **Facilitation** refers to interactions in which one or both species benefit and neither is harmed. Examples include **commensalism** and **mutualism**. A **symbiosis** is a close association in which species are living together.

Communities are characterized by their **structure** (number and size of populations and their interactions) and **dynamics** (how the community structure changes). **Keystone species** have a great influence on community structure, which is disproportionately large relative to their abundance. **Succession** describes how community composition changes over time following a disturbance.

Attribution

Melissa Ha (CC-BY-NC)

This page titled 3.3.4: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **6.4: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

3.4: Ecosystems

Chapter Hook

In semi-urban areas beavers (*Castor canadensis*) are considered a nuisance as their dams block drainage pipes and cause flooding. However, in wild areas beavers are one of the most important species. Beaver dams are known for altering the flow of water. However, these dams not only slow down the movement of water, but they spread and store water in a way that is much more efficient than human-made dams. In addition, they slow and spread sediment and nutrients as they move through watersheds. What this does is create a mosaic of habitats across a landscape, both aquatic and terrestrial. More habitats lead to more species, both for plants and animals. In this way, beavers create ecosystems, one stick at a time.



Figure 3.4.*a* American beaver. Image by NeexPix (Public domain)

Attribution

Rachel Schleiger (CC-BY-NC)

3.4.1: Ecosystem Types and Dynamics

3.4.2: Matter

3.4.3: Biogeochemical Cycles

3.4.4: Soils

3.4.5: Soil Degradation

3.4.6: Data Dive- Beaver Impacts on Wetlands

3.4.7: Review

Thumbnail image - This sage thrasher’s diet, like that of almost all organisms, depends on photosynthesis.

This page titled 3.4: Ecosystems is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

3.4.1: Ecosystem Types and Dynamics

An **ecosystem** is a community of organisms (**biotic components**) and their **abiotic** (non-living) environment. Ecosystems can be small, such as the tide pools found near the rocky shores of many oceans, or large, such as those found in the tropical rainforest of the Amazon in Brazil (figure 3.4.1.*a*).



(a) (b)

Figure 3.4.1.*a*: A (a) tidal pool ecosystem in Matinicus Island, Maine, is a small ecosystem, while the (b) Amazon rainforest in Brazil is a large ecosystem. (credit a: modification of work by Jim Kuhn; credit b: modification of work by Ivan Mlinaric)

Ecosystem Categories

There are three broad categories of ecosystems based on their general environment: freshwater, marine, and terrestrial. Within these three categories are individual ecosystem types based on the environmental habitat and organisms present.

Freshwater ecosystems are the least common, occurring on only 1.8 percent of Earth's surface. These systems comprise lakes, rivers, streams, and springs; they are quite diverse and support a variety of animals, plants, fungi, protists and prokaryotes.

Marine ecosystems are the most common, comprising 75 percent of Earth's surface and consisting of three basic types: shallow ocean, deep ocean water, and deep ocean bottom. Shallow ocean ecosystems include extremely biodiverse coral reef ecosystems. Small photosynthetic organisms suspended in ocean waters, collectively known as **phytoplankton,** perform 40 percent of all photosynthesis on Earth. Deep ocean bottom ecosystems contain a wide variety of marine organisms. These ecosystems are so deep that light is unable to reach them. Freshwater and marine ecosystems are found in **aquatic biomes**, which are discussed in the Biomes chapter.

Terrestrial ecosystems, also known for their diversity, are grouped into large categories called biomes. A **biome** is a large-scale community of organisms, primarily defined on land by the dominant plant types that exist in geographic regions of the planet with similar climatic conditions. Examples of biomes include tropical rainforests, savannas, deserts, grasslands, temperate forests, and tundras. Grouping these ecosystems into just a few biome categories obscures the great diversity of the individual ecosystems within them. For example, the saguaro cacti (*Carnegiea gigantean*) and other plant life in the Sonoran Desert, in the United States, are relatively diverse compared with the desolate rocky desert of Boa Vista, an island off the coast of Western Africa (figure 3.4.1.*b*).



(a)



(b)

Figure 3.4.1.*b*: Desert ecosystems, like all ecosystems, can vary greatly. The desert in (a) Saguaro National Park, Arizona, has abundant plant life, while the rocky desert of (b) Boa Vista island, Cape Verde, Africa, is devoid of plant life. (credit a: modification of work by Jay Galvin; credit b: modification of work by Ingo Wölbern)

Ecosystem Dynamics

Ecosystems are complex with many interacting parts. They are routinely exposed to various **disturbances**, or changes in the environment that affect their compositions: yearly variations in rainfall and temperature and the slower processes of plant growth, which may take several years. Many of these disturbances are a result of natural processes. For example, when lightning causes a forest fire and destroys part of a forest ecosystem, the ground is eventually populated by grasses, then by bushes and shrubs, and later by mature trees, restoring the forest to its former state. The impact of environmental disturbances caused by human activities is as important as the changes caused by natural processes. Human agricultural practices, air pollution, acid rain, global deforestation, overfishing, eutrophication, oil spills, and illegal dumping on land and into the ocean are all issues of concern to conservationists.

Equilibrium is the steady state of an ecosystem where all organisms are in balance with their environment and with each other. In ecology, two parameters are used to measure changes in ecosystems: resistance and resilience. The ability of an ecosystem to remain at equilibrium in spite of disturbances is called **resistance**. The speed at which an ecosystem recovers equilibrium after being disturbed, called its **resilience**. Ecosystem resistance and resilience are especially important when considering human impact. The nature of an ecosystem may change to such a degree that it can lose its resilience entirely. This process can lead to the complete destruction or irreversible altering of the ecosystem.

Foundation Species

Foundation species are considered the “base” or “bedrock” of an ecosystem, having the greatest influence on its overall structure. They are often primary producers, and they are typically an abundant organism. For example, kelp, a species of brown algae, is a foundation species that forms the basis of the kelp forests off the coast of California.

Foundation species may physically modify the environment to produce and maintain habitats that benefit the other organisms that use them. Examples include the kelp described above or tree species found in a forest. The photosynthetic corals of the coral reef also provide structure by physically modifying the environment (figure 3.4.1.*c*). The calcium carbonate deposits of living and dead coral make up most of the reef structure, which protects many other species from waves and ocean currents.

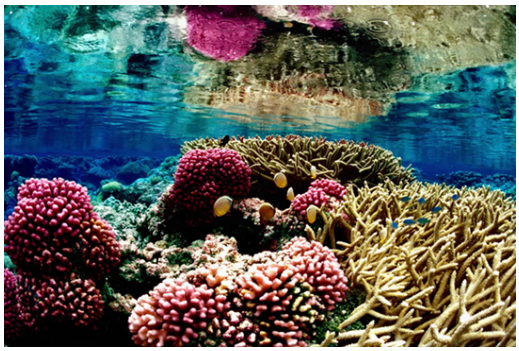


Figure 3.4.1.*c*: Coral is the foundation species of coral reef ecosystems. (credit: Jim E. Maragos, USFWS)

Attributions

Modified by Melissa Ha from the following sources:

- Community Ecology from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Ecology of Ecosystems from *General Biology* by OpenStax (licensed under CC-BY)

This page titled 3.4.1: Ecosystem Types and Dynamics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **7.1: Ecosystem Types and Dynamics** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.4.2: Matter

Matter is the "stuff" found in ecosystems. Technically, **matter** is defined as anything that occupies space or has mass. **Mass** is resistance to acceleration. Put more simply, mass is similar to weight, but weight accounts for acceleration due to gravity. Matter moves between biotic and abiotic ecosystem components through **biogeochemical cycles**. Fully understanding these cycles, requires a background in the particles that comprise matter, atoms.

Atoms and Molecules

Elements are substances that cannot be broken down or transformed chemically into other substances (figure 3.4.2.*a*). A total of 118 elements have been defined; however, only 92 occur naturally and fewer than 30 are found in organisms. The remaining 26 elements are unstable and therefore do not exist for very long or are theoretical and have yet to be detected. Each element is designated by its chemical symbol (such as H, N, O, C, and Na), and possesses unique properties. These unique properties allow elements to combine and to bond with each other in specific ways.

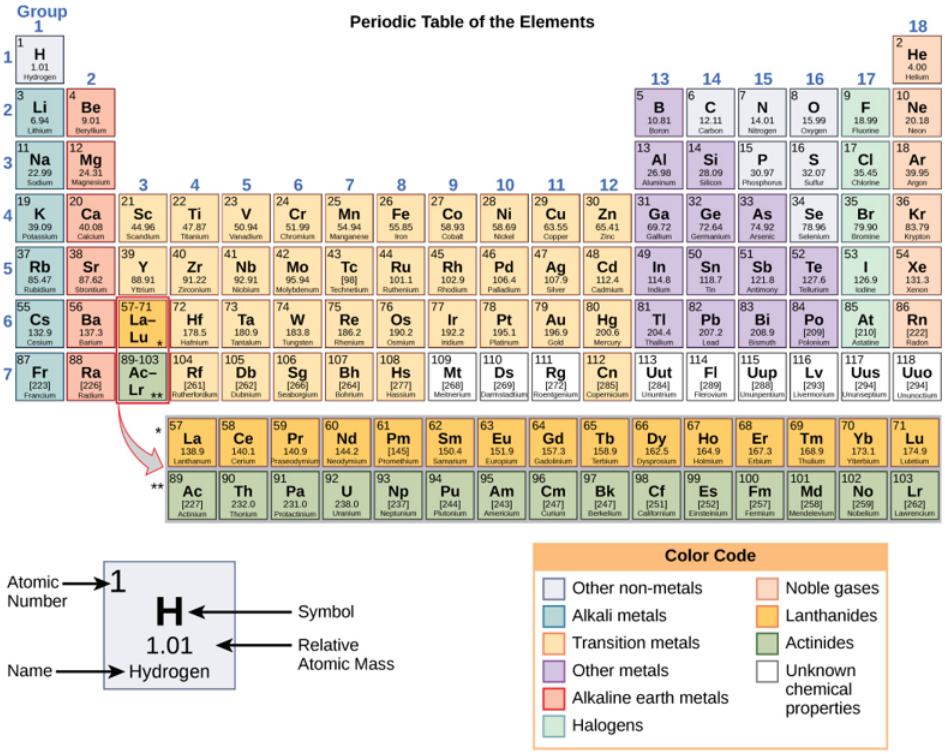


Figure 3.4.2.*a*: The periodic table of elements organizes chemical elements based on their properties. Each element is represented by a one- or two-letter symbol. Additionally, the atomic number, name, and atomic mass are listed for each element. Click here for an accessible periodic table of elements. Image by OpenStax (CC-BY)

An **atom** is the smallest component of an element that retains all of the chemical properties of that element. For example, one hydrogen atom has all of the properties of the element hydrogen, such as it exists as a gas at room temperature and it bonds with oxygen to create a water molecule. Hydrogen atoms cannot be broken down into anything smaller while still retaining the properties of hydrogen. If a hydrogen atom were broken down into smaller particles, it would no longer have the properties of hydrogen.

All atoms contain protons, electrons, and neutrons (figure 3.4.2.*b*). The only exception is hydrogen (H), which is typically only made of one proton and one electron. A **proton** is a positively charged particle that resides in the nucleus (the core of the atom) of an atom and has a mass of 1 and a charge of +1. An **electron** is a negatively charged particle that travels in the space around the nucleus. In other words, it resides outside of the nucleus. It has a negligible mass and has a charge of –1. **Neutrons**, like protons, reside in the nucleus of an atom. They have a mass of 1 and no charge. The positive (protons) and negative (electrons) charges balance each other in a neutral atom, which has a net zero charge.

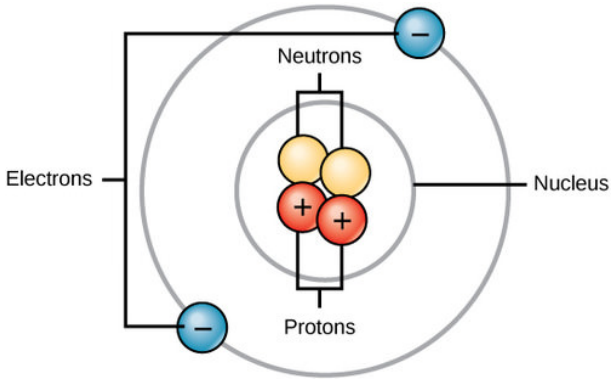


Figure 3.4.2.*b*: Atoms are comprised of protons (red with a plus sign, +) and neutrons (yellow) located within the nucleus, and electrons (blue with a minus sign, −)surrounding the nucleus. The nucleus of an atom should not be confused with the nucleus of a cell.

Each element contains a different number of protons and neutrons, giving it its own atomic number and mass number. The **atomic number** of an element is equal to the number of protons that element contains. The **mass number** is the number of protons plus the number of neutrons of that element. Therefore, it is possible to determine the number of neutrons by subtracting the atomic number from the mass number.

At the most basic level, all organisms are made of a combination of elements. They contain atoms that combine together to form **molecules**. Molecules are chemicals made from two or more atoms bonded together. Some molecules are very simple, like O_2, which is comprised of just two oxygen atoms. Some molecules used by organisms, such as DNA, are made of many millions of atoms.Molecules can interact to form **cells**, the structural and functional units of life. In multicellular organisms, such as animals, these cells combine to form tissues, which make up organs. These combinations continue until entire multicellular organisms are formed.

Chemical Bonds

How elements interact with one another depends on the number of electrons and how they are arranged. When an atom does not contain equal numbers of protons and electrons it is called an **ion**. Because the number of electrons does not equal the number of protons, each ion has a net **charge**. For example, if sodium loses an electron, it now has 11 protons and only 10 electrons, leaving it with an overall charge of +1. Positive ions are formed by losing electrons and are called **cations**. Negative ions are formed by gaining electrons and are called **anions**. Elemental anionic names are changed to end in -ide. As an example, when chlorine becomes an ion it is referred to as chloride.

Ionic and covalent bonds are strong bonds formed between two atoms. These bonds hold atoms together in a relatively stable state. **Ionic bonds** are formed between two oppositely charged ions (an anion and a cation). Because positive and negative charges attract, these ions are held together much like two oppositely charged magnets would stick together. **Covalent bonds** form when electrons are shared between two atoms. Each atom shares one of their electrons, which then orbits the nuclei of both atoms, holding the two atoms together. Covalent bonds are the strongest and most common form of chemical bond in organisms. Unlike most ionic bonds, covalent bonds do not dissociate in water. **Hydrogen bonds** form when molecules have an uneven distribution of electrons and thus have partially positive and partially negative ends. They are thus attracted to each other (figure 3.4.2.*c*). Technically, hydrogen bonds only occur between hydrogen and either oxygen (O), nitrogen (N), or fluorine (F). Sometimes hydrogen bonds connect different parts or large molecules, as is the case in DNA and proteins. Hydrogen bonds are weaker than ionic and covalent bonds and can break easily. (Note that hydrogen bonds are among the strongest of intermolecular forces, those that occur **between** molecules, however.)

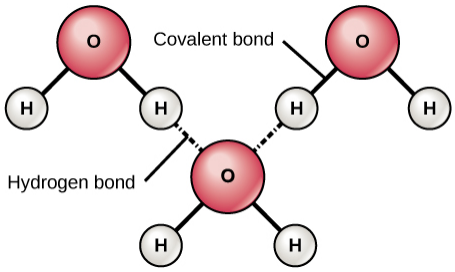


Figure 3.4.2.*c*: Hydrogen bonds form between slightly positive and slightly negative charges of certain molecules, such as water. Each water molecule consists of an oxygen atom (O) connected to two hydrogen atoms (H) by covalent bonds (solid lines). Dashed lines represent hydrogen bonds between water molecules.

Biological Macromolecules

Organisms contain large, organic molecules called **biological macromolecules**. **Organic molecules** are those that contain carbon covalently bonded to hydrogen. (In contrast, **inorganic molecules** lack carbon bonded to hydrogen and are often simpler than organic molecules.) In addition, they may contain oxygen, nitrogen, phosphorus, sulfur, and additional elements.There are four major classes of biological macromolecules: **carbohydrates**, **lipids**, **proteins**, and **nucleic acids.** Each is an important component of the cell and performs a wide array of functions.

It is often said that life is “carbon-based”. This means that carbon atoms, bonded to other carbon atoms or other elements, form the fundamental components of many of the molecules found uniquely in living things. Other elements play important roles in biological molecules, but carbon certainly qualifies as the “foundation” element for molecules in living things. It is the bonding properties of carbon atoms that are responsible for its important role. Carbon can form four covalent bonds with other atoms or molecules. The simplest organic carbon molecule is methane (CH_4), in which four hydrogen atoms bind to a carbon atom (figure 3.4.2.*d*).

Carbohydrates include what are commonly referred to as simple sugars, like glucose, and complex carbohydrates such as starch. While many types of carbohydrates are used for energy, some are used for structure by most organisms, including plants and animals. For example, cellulose is a complex carbohydrate that adds rigidity and strength to the outer layer of plant cells (the cell walls).

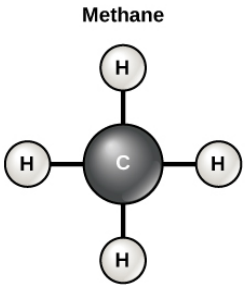


Figure 3.4.2.*d*: Carbon can form four covalent bonds to create an organic molecule. The simplest carbon molecule is methane (CH_4), depicted here.

Lipids include a diverse group of compounds that are united by a common feature: lipids are insoluble in water. Lipids perform many different functions in a cell. Cells store energy for long-term use in the form of lipids called fats. Lipids also provide insulation from the environment for plants and animals. For example, they help keep aquatic birds and mammals dry because of their water-repelling nature. Lipids are also the building blocks of many hormones and are an important constituent of the membranes that surround cells and form many of their internal structures.

Proteins are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules. They are all polymers of amino acids. The functions of proteins are very diverse because there are 20 different chemically distinct amino acids that form long chains, and the amino acids can be in any order. Proteins can function in facilitated chemical reactions in organisms, such as photosynthesis, transmitting messages as hormones, causing muscles to contract, and much more.

Nucleic acids are very large molecules that are important to the continuity of life. They carry the genetic blueprint of a cell and thus the instructions for its functionality. The two main types of nucleic acids are deoxyribonucleic acid (**DNA**) and ribonucleic acid (RNA). DNA is the genetic material found in all organisms, ranging from single-celled bacteria to multicellular mammals. The other type of nucleic acid, **RNA**, is mostly involved in protein synthesis. DNA and RNA are made up of small building blocks known as **nucleotides**. DNA has a beautiful double-helical structure (Figure 3.4.2.e).

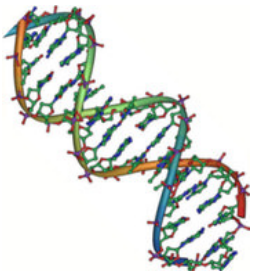


Figure 3.4.2.e: The double-helix model shows DNA as two parallel strands of intertwining molecules.

Attribution

Modified by Melissa Ha from Matter from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 3.4.2: Matter is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **7.2: Matter** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.4.3: Biogeochemical Cycles

Biogeochemical cycles, also known as nutrient cycles, describe the movement of chemical elements through different media, such as the atmosphere, soil, rocks, bodies of water, and organisms. Biogeochemical cycles keep essential elements available to plants and other organisms.

Energy flows directionally through ecosystems, entering as sunlight (or inorganic molecules for chemoautotrophs) and leaving as heat during energy transformation between trophic levels. Rather than flowing through an ecosystem, the matter that makes up organisms is conserved and recycled. The **law of conservation of mass** states that matter is neither created nor destroyed. For example, after a chemical reaction, the mass of the products (ending molecules) will be the same as the mass of the reactants (starting molecules). The same is true in an ecosystem. Matter moves through different media, and atoms may react to form new molecules, but the amount of matter remains constant.

The biogeochemical cycles of four elements—carbon, nitrogen, phosphorus, and sulfur—are discussed below. The cycling of these elements is interconnected with the water cycle. For example, the movement of water is critical for the leaching of sulfur and phosphorus into rivers, lakes, and oceans. Today, **anthropogenic** (human) activities are altering all major ecosystems and the biogeochemical cycles they drive.

The Carbon Cycle

Carbon is the basic building block of all organic materials, and therefore, of living organisms. The carbon cycle is actually comprised of several interconnected cycles: one dealing with rapid carbon exchange among living organisms and the other dealing with the long-term cycling of carbon through geologic processes (figure $3.4.3.a$). The overall effect is that carbon is constantly recycled in the dynamic processes taking place in the atmosphere, at the surface and in the crust of the earth. The vast majority of carbon resides as inorganic minerals in crustal rocks. Other **reservoirs** of carbon, places where carbon accumulates, include the oceans and atmosphere. Some of the carbon atoms in your body today may long ago have resided in a dinosaur's body, or perhaps were once buried deep in the Earth's crust as carbonate rock minerals.

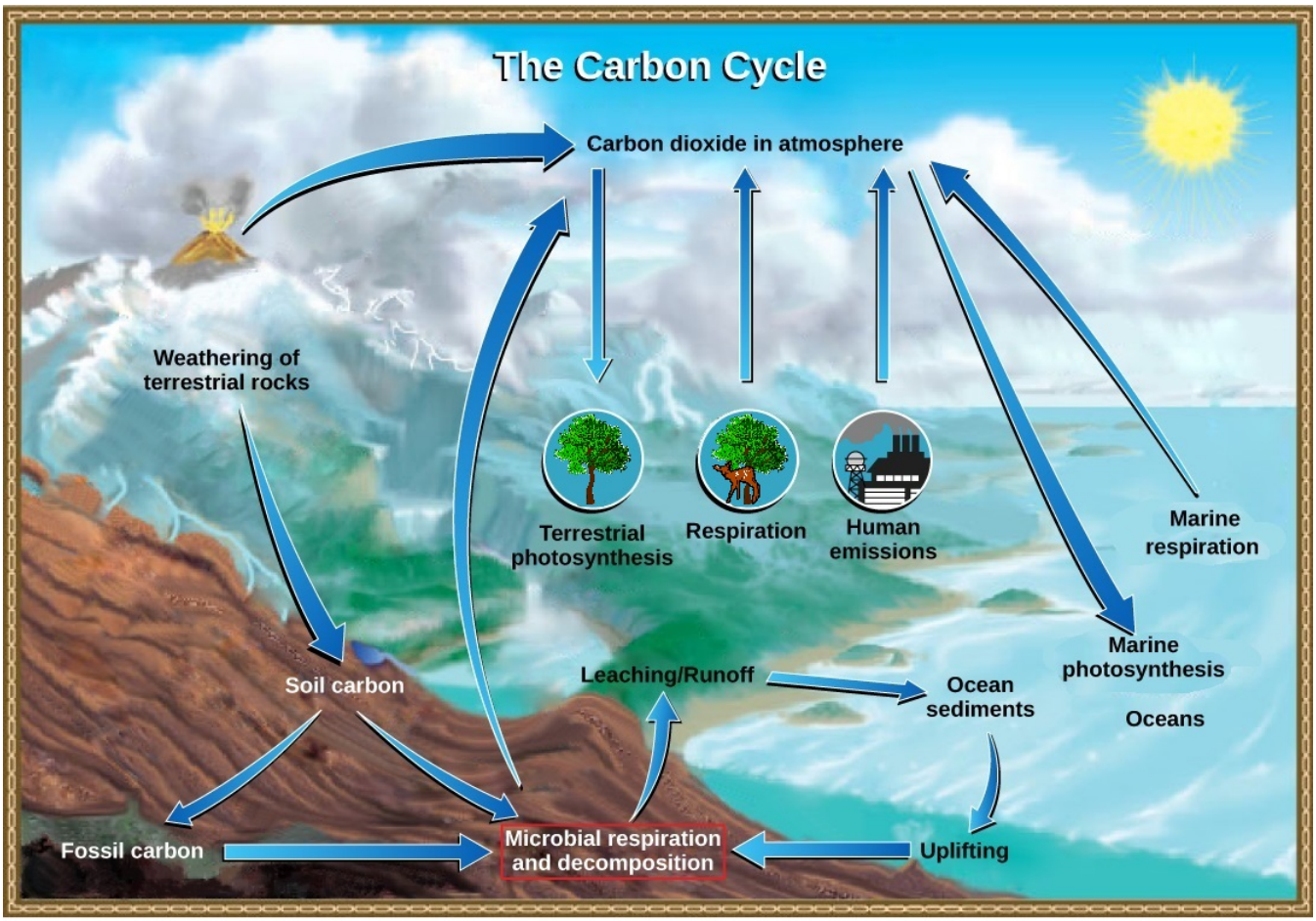


Figure 3.4.3.a: Carbon dioxide in the atmosphere is converted to organic carbon through photosynthesis by terrestrial organisms (like trees) and marine organisms (like algae). Respiration by terrestrial organisms (like trees and deer) and marine organisms (like algae and fish) release carbon dioxide back into the atmosphere. Additionally, microbes that decompose dead organisms release carbon dioxide through respiration. Weathering of terrestrial rocks also brings carbon into the soil. Carbon in the soil enters the water through leaching and runoff. It can accumulate into ocean sediments and reenter land through uplifting. Long-term storage of organic carbon occurs when matter from living organisms is buried deep underground and becomes fossilized. Volcanic activity and, more recently, human emissions stored carbon back into the carbon cycle. Modified from John M. Evans and Howard Perlman, USGS using tree and deer (both public domain).

Carbon Cycles Slowly between Land and the Ocean

On land, carbon is stored in soil as organic carbon in the form of decomposing organisms or terrestrial rocks. Decomposed plants and algae are sometimes buried and compressed between layers of sediments. After millions of years fossil fuels such as coal, oil, and natural gas are formed. The **weathering** of terrestrial rock and minerals release carbon into the soil.

Carbon-containing compounds in the soil can be washed into bodies of water through **leaching.** This water eventually enters the ocean. Atmospheric carbon dioxide also dissolves in the ocean, reacting with water molecules to form carbonate ions (CO_3^{2-}). Some of these ions combine with calcium ions in the seawater to form calcium carbonate ($CaCO_3$), a major component of the shells of marine organisms. These organisms eventually die and their shells form sediments on the ocean floor. Over geologic time, the calcium carbonate forms limestone, which comprises the largest carbon reservoir on Earth.

Carbonate also precipitates in sediments, forming carbonate rocks, such as limestone. Carbon sediments from the ocean floor are taken deep within Earth by the process of **subduction**: the movement of one tectonic plate beneath another. The ocean sediments are subducted by the actions of **plate tectonics**, melted and then returned to the surface during volcanic activity. Plate tectonics can also cause **uplifting**, returning ocean sediments to land.

Carbon Cycles Quickly between Organisms and the Atmosphere

Carbon dioxide is converted into glucose, an energy-rich organic molecule through **photosynthesis** by plants, algae, and some bacteria (figure 3.4.3.*b*). They can then produce other organic molecules like complex carbohydrates (such as starch), proteins and lipids, which animals can eat. Most terrestrial autotrophs obtain their carbon dioxide directly from the atmosphere, while marine autotrophs acquire it in the dissolved form (bicarbonate, HCO_3^-).



Figure 3.4.3.*b*: (a) Plants, (b) algae, and (c) certain bacteria, called cyanobacteria, are can carry out photosynthesis. Algae can grow over enormous areas in water, at times completely covering the surface. (credit a: Steve Hillebrand, U.S. Fish and Wildlife Service; credit b: "eutrophication&hypoxia"/Flickr; credit c: NASA; scale-bar data from Matt Russell)

Plants, animals, and other organisms break down these organic molecules during the process of **aerobic cellular respiration**, which consumes oxygen and releases energy, water and carbon dioxide. Carbon dioxide is returned to the atmosphere during gaseous exchange. Another process by which organic material is recycled is the decomposition of dead organisms. During this process, bacteria and fungi break down the complex organic compounds. Decomposers may do respiration, releasing carbon dioxide, or other processes that release methane (CH_4).

Photosynthesis and respiration are actually reciprocal to one another with regard to the cycling of carbon: photosynthesis removes carbon dioxide from the atmosphere and respiration returns it (figure 3.4.3.*c*). A significant disruption of one process can therefore affect the amount of carbon dioxide in the atmosphere.

Photosynthesis Equation							
Carbon dioxide	+	Water	SUNLIGHT →	Sugar	+	Oxygen	
$6CO_2$		$6H_2O$		$C_6H_{12}O_6$		$6O_2$	

Figure 3.4.3.*c*: This equation means that six molecules of carbon dioxide (CO_2) combine with six molecules of water (H_2O) in the presence of sunlight. This produces one molecule of glucose ($C_6H_{12}O_6$) and six molecules of oxygen (O_2).

Cellular respiration is only one process that releases carbon dioxide. Physical processes, such as the eruption of volcanoes and release from **hydrothermal vents** (openings in the ocean floor) add carbon dioxide to the atmosphere. Additionally, the **combustion** of wood and fossil fuels releases carbon dioxide. The level of carbon dioxide in the atmosphere is greatly influenced by the reservoir of carbon in the oceans. The exchange of carbon between the atmosphere and water reservoirs influences how much carbon is found in each.

Importance of the Carbon Cycle

The carbon cycle is crucially important to the biosphere. If not for the recycling processes, carbon might long ago have become completely sequestered in crustal rocks and sediments, and life would no longer exist (figure 3.4.3.*e*). Photosynthesis not only makes energy and carbon available to higher trophic levels, but it also releases gaseous oxygen (O_2). Gaseous oxygen is necessary for cellular respiration to occur. Photosynthetic bacteria were likely the first organisms to perform photosynthesis, dating back 2-3 billion years ago. Thanks to their activity, and a diversity of present-day photosynthesizing organisms, Earth's atmosphere is currently about 21% O_2. Also, this O_2 is vital for the creation of the ozone layer, which protects life from harmful ultraviolet radiation emitted by the sun. Ozone (O_3) is created from the breakdown and reassembly of O_2.



Figure 3.4.3.*e*: Decomposers will break down the organic compounds in this fallen tree at Cliffs of the Neuse State Park in Wayne County, North Carolina, releasing carbon dioxide into the atmosphere. Decomposition ensures that carbon dioxide will be available in the atmosphere for photosynthetic organisms, which then provide carbon for consumers. Image by Gerry Dincher (CC-BY-SA).

The global carbon cycle contributes substantially to the provisioning ecosystem services upon which humans depend. We harvest approximately 25% of the total plant biomass that is produced each year on the land surface to supply food, fuel wood and fiber from croplands, pastures and forests. In addition, the global carbon cycle plays a key role in regulating ecosystem services because it significantly influences climate via its effects on atmospheric CO_2 concentrations.

Human Alteration of the Carbon Cycle

Atmospheric CO_2 concentration increased from 280 parts per million (ppm) to 413 ppm between the start of industrial revolution in the late eighteenth century and 2020. This reflected a new flux in the global carbon cycle—anthropogenic CO_2 emissions—where humans release CO_2 into the atmosphere by burning fossil fuels and changing land use. Fossil fuel burning takes carbon from coal, gas, and oil reserves, where it would be otherwise stored on very long time scales, and introduces it into the active carbon cycle. Land use change releases carbon from soil and plant biomass pools into the atmosphere, particularly through the process of deforestation for wood extraction or conversion of land to agriculture. In 2018, the additional flux of carbon into the atmosphere from anthropogenic sources was estimated to be 36.6 gigatons of carbon (GtC = 1 billion tons of carbon)—a significant disturbance to the natural carbon cycle that had been in balance for several thousand years previously. High levels of carbon dioxide in the atmosphere cause warming that results in climate change. (See Threats to Biodiversity and Climate Change for more details.)

The Nitrogen Cycle

All organisms require nitrogen because it is an important component of nucleic acids, proteins, and other organic molecules. Getting nitrogen into living organisms is difficult. Plants and algae are not equipped to incorporate nitrogen from the atmosphere (where it exists as tightly bonded, triple covalent N_2) although this molecule comprises approximately 78 percent of the atmosphere. Because most of the nitrogen is stored in the atmosphere, the atmosphere is considered a reservoir of nitrogen.

The nitrogen molecule (N_2) is quite inert. To break it apart so that its atoms can combine with other atoms requires the input of substantial amounts of energy. **Nitrogen fixation** is the process of converting nitrogen gas into ammonia (NH_3), which spontaneously becomes ammonium (NH_4^+). Ammonium is found in bodies of water and in the soil (figure 3.4.3.*f*).

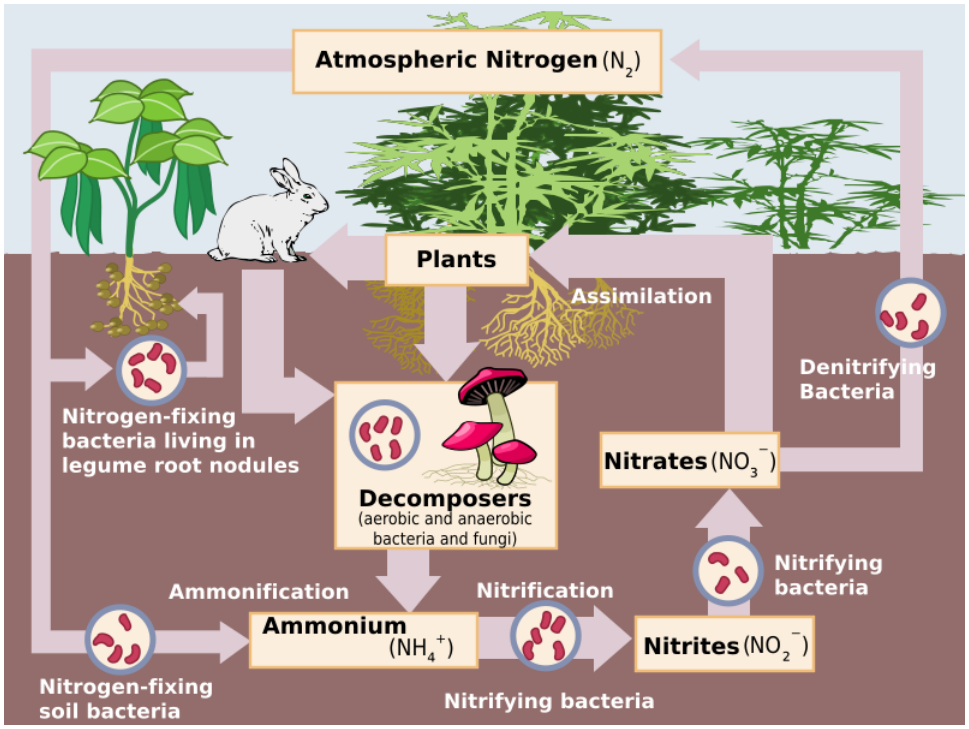


Figure 3.4.3.*f*: In the nitrogen cycle, nitrogen-fixing bacteria in the soil or legume root nodules convert nitrogen gas (N_2) from the atmosphere to ammonium (NH_4^+). Nitrification occurs when bacteria convert ammonium to nitrites (NO_2^-) and then to nitrates (NO_3^-). Nitrates re-enter the atmosphere as nitrogen gas through denitrification by bacteria. Plants assimilate ammonium and nitrates, producing organic nitrogen, which is available to consumers. Decomposers, including aerobic and anaerobic bacteria and fungi, break down organic nitrogen and release ammonium through ammonification. (credit: "Nitrogen cycle" by Johann Dréo & Raeky is licensed under CC BY-SA 3.0)

Three processes are responsible for most of the nitrogen fixation in the biosphere. The first is **atmospheric fixation** by lightning. The enormous energy of lightning breaks nitrogen molecules and enables their atoms to combine with oxygen in the air forming nitrogen oxides. These dissolve in rain, forming nitrates, that are carried to the earth. Atmospheric nitrogen fixation probably contributes some 5-8% of the total nitrogen fixed. The second process is **industrial fixation.** Under great pressure, at a temperature of 600°C (1112°F), and with the use of a **catalyst** (which facilitates chemical reactions), atmospheric nitrogen and hydrogen can be combined to form ammonia (NH_3). Ammonia can be used directly as fertilizer, but most of it is further processed to urea and ammonium nitrate (NH_4NO_3).

The third process is **biological fixation** by certain free-living or symbiotic bacteria. Some form a symbiotic relationship with plants in the legume family, which includes beans, peas, soybeans, alfalfa, and clovers (figure 3.4.3.*g*). Some nitrogen-fixing bacteria even establish symbiotic relationships with animals, e.g., termites and "shipworms" (wood-eating bivalves). Nitrogen-fixing cyanobacteria are essential to maintaining the fertility of semi-aquatic environments like rice paddies. Although the first stable product of the process is ammonia, this is quickly incorporated into protein and other organic nitrogen compounds.



Figure 3.4.3.*g*: Nitrogen-fixing bacteria live in the spherical nodules of this soybean root. Image by United Soybean Board (CC-BY).

Ammonium is converted by bacteria and archaea into nitrites (NO_2^-) and then nitrates (NO_3^-) through the process of **nitrification**. Like ammonium, nitrites and nitrates are found in water and the soil. Some nitrates are converted back into nitrogen gas, which is released into the atmosphere. The process, called **denitrification**, is conducted by bacteria.

Plants and other producers directly use ammonium and nitrates to make organic molecules through the process of **assimilation**. This nitrogen is now available to consumers. Organic nitrogen is especially important to the study of ecosystem dynamics because many processes, such as primary production, are limited by the available supply of nitrogen.

Consumers excrete organic nitrogen compounds that return to the environment. Additionally dead organisms at each trophic level contain organic nitrogen. Microorganisms, such as bacteria and fungi, decompose these wastes and dead tissues, ultimately producing ammonium through the process of **ammonification.**

In marine ecosystems, nitrogen compounds created by bacteria, or through decomposition, collects in ocean floor sediments. It can then be moved to land in geologic time by uplift of Earth's crust and thereby incorporated into terrestrial rock. Although the movement of nitrogen from rock directly into living systems has been traditionally seen as insignificant compared with nitrogen fixed from the atmosphere, a recent study showed that this process may indeed be significant and should be included in any study of the global nitrogen cycle.

Human activity can alter the nitrogen cycle by two primary means: the combustion of fossil fuels, which releases different **nitrogen oxides** into the atmosphere, and by the use of artificial fertilizers in agriculture. Atmospheric nitrogen (other than N_2) is associated with several effects on Earth's ecosystems. Nitrogen oxides (HNO_3) can react in the atmosphere to form nitric acid, a form of **acid deposition**, also known as acid rain. Acid deposition damages healthy trees, destroys aquatic systems and erodes building materials such as marble and limestone. Like carbon dioxide, nitrous oxide (N_2O) causes warming resulting in climate change.

Humans are primarily dependent on the nitrogen cycle as a supporting ecosystem service for crop and forest productivity. Nitrogen fertilizers are added to enhance the growth of many crops and plantations (figure 3.4.3.*h*). The enhanced use of fertilizers in agriculture was a key feature of the green revolution that boosted global crop yields in the 1970s. The industrial production of nitrogen-rich fertilizers has increased substantially over time and now matches more than half of the input to the land from biological nitrogen fixation (90 megatons = 1 million tons of nitrogen each year). If the nitrogen fixation from legume crops is included, then the anthropogenic flux of nitrogen from the atmosphere to the land exceeds natural fluxes to the land. Fertilizers are washed into lakes, streams, and rivers by surface runoff, resulting in saltwater and freshwater **eutrophication**, a process whereby nutrient runoff causes the overgrowth of algae, the depletion of oxygen, and death of aquatic fauna.



Figure 3.4.3.*h*: Fertilizer containing nitrogen is conventionally applied at large scales in agriculture. Image by Bob Nichols, USDA Natural Resources Conservation Service (public domain).

The Phosphorus Cycle

Several forms of nitrogen (nitrogen gas, ammnoium, nitrates, etc.) were involved in the nitrogen cycle, but phosphorus remains primarily in the form of the phosphate ion (PO_4^{3-}). Also in contrast to the nitrogen cycle, there is no form of phosphorus in the atmosphere. Phosphorus is used to make nucleic acids and the phospholipids that comprise biological membranes.

Rocks are a reservoir for phosphorus, and these rocks have their origins in the ocean. Phosphate-containing ocean sediments form primarily from the bodies of ocean organisms and from their excretions. However, volcanic ash, aerosols, and mineral dust may also be significant phosphate sources. This sediment then is moved to land over geologic time by the uplifting of Earth's surface (figure 3.4.3.*i*). The movement of phosphate from the ocean to the land and through the soil is extremely slow, with the average phosphate ion having an oceanic residence time between 20,000 and 100,000 years.

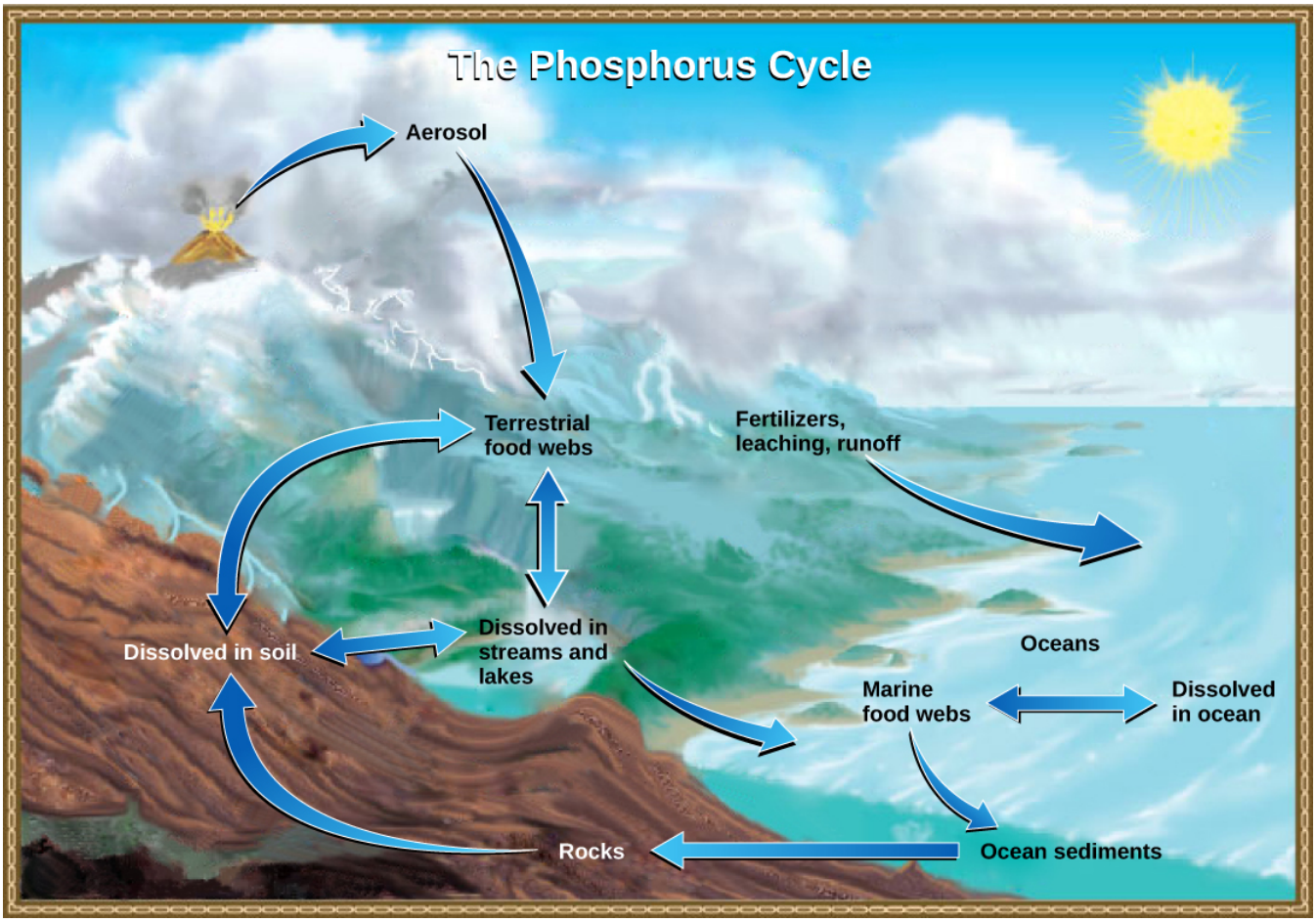


Figure 3.4.3.*i*: In nature, phosphorus exists as the phosphate ion (PO_4^{3-}). Phosphate enters the atmosphere from volcanic aerosols, which precipitate to Earth. Weathering of rocks also releases phosphate into the soil and water, where it becomes available to terrestrial food webs. Some of the phosphate from terrestrial food webs dissolves in streams and lakes, and the remainder enters the soil. Phosphate enters the ocean via surface runoff, groundwater flow, and river flow, where it becomes dissolved in ocean water or enters marine food webs. Some phosphate falls to the ocean floor where it becomes sediment. If uplifting occurs, this sediment can return to land. (credit: modification of work by John M. Evans and Howard Perlman, USGS)

Marine birds play a unique role in the phosphorous cycle. These birds take up phosphorous from ocean fish. Their droppings on land (**guano**) contain high levels of phosphorous and are sometimes mined for commercial use. A 2020 study estimated that the **ecosystem services** (natural processes and products that benefit humans) provided by guano are worth $470 million per year.

Weathering of rocks releases phosphates into the soil and bodies of water. Plants can assimilate phosphates in the soil and incorporate it into organic molecules, making phosphorus available to consumers in terrestrial food webs. Waste and dead organisms are decomposed by fungi and bacteria, releasing phosphates back into the soil. Some phosphate is leached from the soil, entering into rivers, lakes, and the ocean. Primary producers in aquatic food webs, such as algae and photosynthetic bacteria, assimilate phosphate, and organic phosphate is thus available to consumers in aquatic food webs. Similar to terrestrial food webs, phosphorus is reciprocally exchanged between phosphate dissolved in the ocean and organic phosphorus in marine organisms.

The movement of phosphorus from rock to living organisms is normally a very slow process, but some human activities speed up the process. Phosphate-bearing rock is often mined for use in the manufacture of fertilizers and detergents. This commercial production greatly accelerates the phosphorous cycle. In addition, runoff from agricultural land and the release of sewage into water systems can cause a local overload of phosphate. The increased availability of phosphate can cause overgrowth of algae. This reduces the oxygen level, causing eutrophication and the destruction of other aquatic species.

Eutrophication and Dead Zones

Eutrophication occurs when excess phosphorus and nitrogen from fertilizer runoff or sewage causes excessive growth of algae. Algal blooms that block light and therefore kill aquatic plants in rivers, lakes, and seas. The subsequent death and decay of these

organisms depletes dissolved oxygen, which leads to the death of aquatic organisms such as shellfish and fish. This process is responsible for **dead zones**, large areas in lakes and oceans near the mouths of rivers that are periodically depleted of their normal flora and fauna, and for massive fish kills, which often occur during the summer months (figure 3.4.3.*j*). There are more than 500 dead zones worldwide. One of the worst dead zones is off the coast of the United States in the Gulf of Mexico. Fertilizer runoff from the Mississippi River basin created a dead zone, which reached its peak size of 8,776 square miles in 2017. Phosphate and nitrate runoff from fertilizers also negatively affect several lake and bay ecosystems including the Chesapeake Bay in the eastern United States.

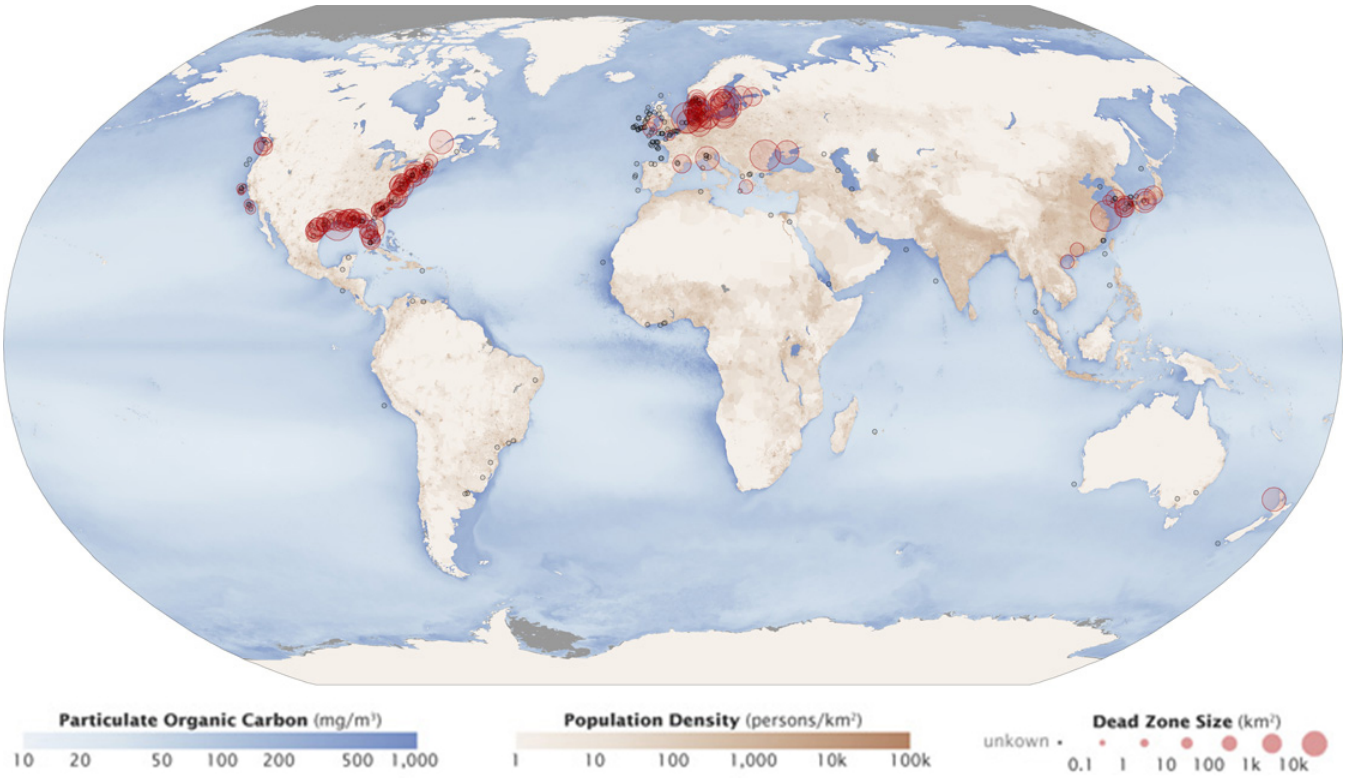
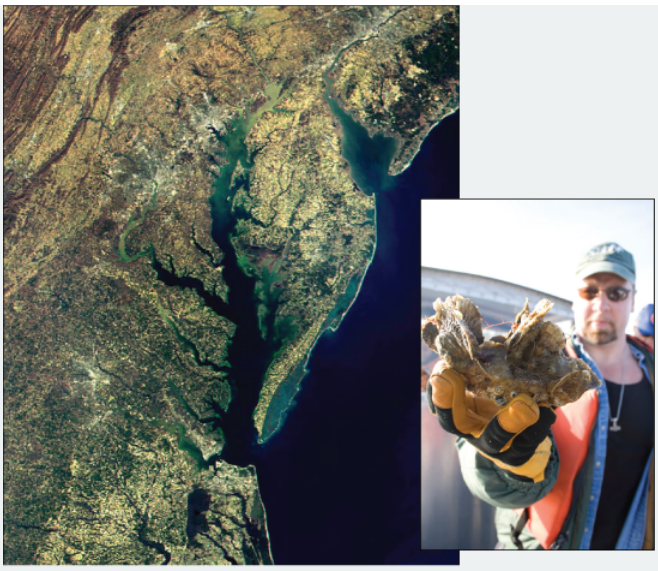


Figure 3.4.3.*j*: Dead zones occur when phosphorus and nitrogen from fertilizers cause excessive growth of microorganisms, which depletes oxygen and kills fauna. This map shows dead zones around the world in 2008. Worldwide, large dead zones are found in coastal areas of high population density. (credit: NASA Earth Observatory)

Everyday Connection: Chesapeake Bay

The Chesapeake Bay has long been valued as one of the most scenic areas on Earth; it is now in distress and is recognized as a declining ecosystem. In the 1970s, the Chesapeake Bay was one of the first ecosystems to have identified dead zones, which continue to kill many fish and bottom-dwelling species, such as clams, oysters, and worms (figure 3.4.3.*k*). Several species have declined in the Chesapeake Bay due to surface water runoff containing excess nutrients from artificial fertilizer used on land. The source of the fertilizers (with high nitrogen and phosphate content) is not limited to agricultural practices. There are many nearby urban areas and more than 150 rivers and streams empty into the bay that are carrying fertilizer runoff from lawns and gardens. Thus, the decline of the Chesapeake Bay is a complex issue and requires the cooperation of industry, agriculture, and everyday homeowners.



(a) (b)

Figure 3.4.3.*k*: This (a) satellite image shows the Chesapeake Bay, an ecosystem affected by phosphate and nitrate runoff. A (b) member of the Army Corps of Engineers holds a clump of oysters being used as a part of the oyster restoration effort in the bay. (credit a: modification of work by NASA/MODIS; credit b: modification of work by U.S. Army)

Of particular interest to conservationists is the oyster population; it is estimated that more than 200,000 acres of oyster reefs existed in the bay in the 1700s, but that number has now declined to only 36,000 acres. Oyster harvesting was once a major industry for Chesapeake Bay, but it declined 88 percent between 1982 and 2007. This decline was due not only to fertilizer runoff and dead zones but also to overexploitation. Oysters require a certain minimum population density because they must be in close proximity to reproduce. Human activity has altered the oyster population and locations, greatly disrupting the ecosystem.

The restoration of the oyster population in the Chesapeake Bay has been ongoing for several years with mixed success. Not only do many people find oysters good to eat, but they also clean up the bay. Oysters are filter feeders, and as they eat, they clean the water around them. In the 1700s, it was estimated that it took only a few days for the oyster population to filter the entire volume of the bay. Today, with changed water conditions, it is estimated that the present population would take nearly a year to do the same job.

Restoration efforts have been ongoing for several years by non-profit organizations, such as the Chesapeake Bay Foundation. The restoration goal is to find a way to increase population density so the oysters can reproduce more efficiently. Many disease-resistant varieties (developed at the Virginia Institute of Marine Science for the College of William and Mary) are now available and have been used in the construction of experimental oyster reefs. Efforts to clean and restore the bay by Virginia and Delaware have been hampered because much of the pollution entering the bay comes from other states, which stresses the need for inter-state cooperation to gain successful restoration.

The new, hearty oyster strains have also spawned a new and economically viable industry—oyster aquaculture—which not only supplies oysters for food and profit, but also has the added benefit of cleaning the bay.

The Sulfur Cycle

Sulfur is an essential element for the molecules of living things. As part of the amino acid cysteine, it is critical to the three-dimensional shape of proteins. As shown in Figure 3.4.3.*l*, sulfur cycles among the oceans, land, and atmosphere. Atmospheric sulfur is found in the form of sulfur dioxide (SO_2), which enters the atmosphere in three ways: first, from the decomposition of organic molecules; second, from volcanic activity and geothermal vents; and, third, from the burning of fossil fuels by humans.

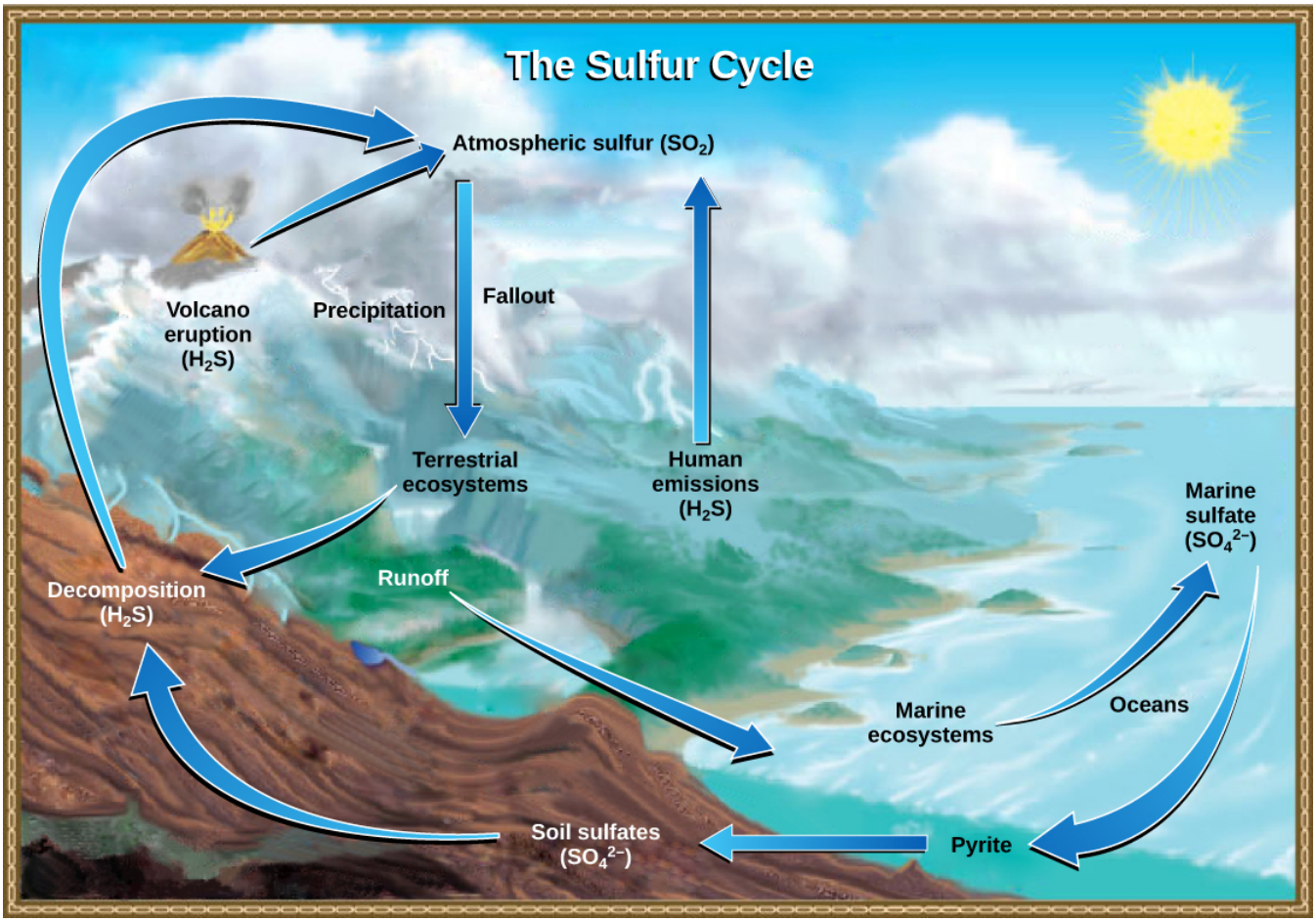


Figure 3.4.3.1: The sulfur cycle. Sulfur dioxide (SO_2) from the atmosphere is dissolved in precipitation as weak sulfuric acid or falls directly to Earth as fallout. This releases sulfates (SO_4^{2-}) into the soil and water. Soil sulfates can be carried as runoff into the water. Marine sulfate can form pyrite, and this can break down to release soil sulfates. Organisms in terrestrial and marine ecosystems assimilate sulfate, adding sulfur to organic molecules, such as proteins (not shown). Decomposition of these organisms returns sulfates to the soil. Microorganisms can convert sulfates to hydrogen sulfide (H_2S) and vice versa. Decomposition, volcanic eruptions, and human activities (including burning fossil fuels) can release hydrogen sulfide (H_2S) or sulfur dioxide into the atmosphere. (credit: modification of work by John M. Evans and Howard Perlman, USGS)

On land, sulfur is deposited in four major ways: precipitation, direct fallout from the atmosphere, rock weathering, and geothermal vents. Atmospheric sulfur is found in the form of sulfur dioxide (SO_2), and as rain falls through the atmosphere, sulfur is dissolved in the form of weak sulfuric acid (H_2SO_4). Sulfur can also fall directly from the atmosphere in a process called **fallout**. Also, as sulfur-containing rocks weather, sulfur is released into the soil. These rocks originate from ocean sediments that are moved to land by the geologic uplifting of ocean sediments. Terrestrial ecosystems can then make use of these soil sulfates (SO_4^{2-}), which enter the food web by being taken up by plant roots. When these plants decompose and die, sulfur is released back into the atmosphere as hydrogen sulfide (H_2S) gas.

Sulfur enters the ocean in runoff from land, from atmospheric fallout, and from hydrothermal vents. Some ecosystems rely on microorganisms using sulfur as a biological energy source (in contrast to ecosystems with photosynthetic producers). This sulfur then supports marine ecosystems in the form of sulfates.

Human activities have played a major role in altering the balance of the global sulfur cycle. The burning of large quantities of fossil fuels, especially from coal, releases sulfur dioxide, which reacts with the atmosphere to form sulfuric acid. Like nitric acid, sulfuric acid contributes to acid deposition.

Suggested Supplementary Reading

Bruckner, M. 2018. The Gulf of Mexico Dead Zone. [Website]

References

Cell Press. (2020, August 6). Researchers hope to save seabirds by calculating the value of their excrement. Retrieved August 7, 2020 from *ScienceDaily*.

Attributions

Modified by Melissa Ha from the following sources:

- Biogeochemical Cycles, Energy, and Energy Enters Ecosystems Through Photosynthesis from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Carbon Cycle and Nitrogen Cycle from *Biology* by John W. Kimball (licensed under CC-BY)
- Energy Flow through Ecosystems from *General Biology* by OpenStax (licensed under CC-BY)
- Soil and Sustainability and Biogeochemical Cycles and the Flow of Energy in the Earth System from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors (licensed under CC-BY). Download for free at CNX.
- Cycling of Matter from *AP Environmental Science* by University of California College Prep, University of California (licensed under CC-BY). Download for free at CNX.
- Nutrient Cycles from *Life Sciences Grade 10* by Siyavula (CC-BY)

This page titled 3.4.3: Biogeochemical Cycles is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **7.3: Biogeochemical Cycles** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

3.4.4: Soils

Soil is the outer loose layer that covers the surface of Earth. Soil quality is a major determinant, along with climate, of plant distribution and growth. Soil quality depends not only on the chemical composition of the soil, but also climate, topography, and organisms living in the soil. In agriculture, the history of the soil, such as the cultivating practices and previous crops, modify the characteristics and fertility of that soil.

Importance of Soil

Soil is important to our society primarily because it provides the foundation of agriculture and forestry. Soil plays a key role in plant growth. Beneficial aspects to plants include providing physical support, water, heat, nutrients, and oxygen. Mineral nutrients from the soil can dissolve in water and then become available to plants. Through their roots, plants absorb water and minerals (e.g., nitrates, phosphates, potassium, copper, zinc). With these and carbon dioxide acquired during photosynthesis, plants produce carbohydrates, proteins, lipids, nucleic acids, and the vitamins, on which consumers depend.

Soil plays a role in nearly all biogeochemical cycles on the Earth's surface. Global cycling of key elements such as carbon (C), nitrogen (N), phosphorous (P), and sulfur (S) all pass through soil. In the hydrologic (water) cycle, soil helps to mediate **infiltration** (percolating) from the surface into the groundwater. Microorganisms living in soil can also be important components of biogeochemical cycles through the action of decomposition and other processes such as nitrogen fixation.

Several elements are considered essential for plant growth. Carbon (C), hydrogen (H), and oxygen (O) are required in large quantities, but they are not absorbed as mineral nutrients from the soil. Plants obtain carbon from carbon dioxide in the atmosphere and hydrogen from the water absorbed by the roots. Oxygen atoms come from carbon dioxide acquired during and gaseous oxygen in the atmosphere (acquire through aerobic cellular respiration) as well as water. Of the mineral nutrients absorbed from the soil, **macronutrients**, including nitrogen (N), potassium (K), calcium (Ca), magnesium (Mg), phosphorus (P), sulfur (S), and silicon (Si) are needed by plants in significant quantities. **Micronutrients** are essential elements that are needed only in small quantities, but can still be limiting to plant growth since these nutrients are not so abundant in nature. Micronutrients include chlorine (Cl), iron (Fe), boron (B), manganese (Mn), sodium (Na), zinc (Zn), copper (Cu), nickel (Ni), and molybdenum (Mo). There are some other elements that tend to aid plant growth but are not absolutely essential.

Although many aspects of soil are beneficial to plants, excessively high levels of trace metals (either naturally occurring or added by humans) or applied herbicides can be toxic to some plants (figure 3.4.4.a).



Figure 3.4.4.a: These native plants are affected by herbicides that were used to kill nearby weeds. Image by Matt Lavin (CC-BY-SA).

Micronutrients and macronutrients are desirable in particular concentrations and can be detrimental to plant growth when concentrations in soil solution are either too low (limiting) or too high (toxicity). Mineral nutrients are useful to plants only if they are in an extractable form in soil solutions, such as a dissolved ion rather than in solid mineral. Many nutrients move through the soil and into the root system as a result of concentration gradients, moving by diffusion from high to low concentrations. However, some nutrients are selectively absorbed by the root membranes, enabling concentrations to become higher inside the plant than in the soil.

An important factor affecting soil fertility is soil **pH** (the negative log of the hydrogen ion concentration). Soil pH is a measure of the acidity or alkalinity of the soil solution. On the pH scale (0 to 14) a value of seven represents a neutral solution; a value less than seven represents an **acidic** solution and a value greater than seven represents an **alkaline** solution (figure 3.4.4.*b*). Soil pH affects the health of microorganisms in the soil and controls the availability of nutrients in the soil solution. Strongly acidic soils (less than 5.5) hinder the growth of bacteria that decompose organic matter in the soil. This results in a buildup of organic matter that has yet to be decomposed, which leaves important nutrients such as nitrogen in forms that are unusable by plants. Soil pH also affects the solubility of nutrient-bearing minerals. This is important because the nutrients must be dissolved in solution for plants to assimilate them through their roots. Most minerals are more soluble in slightly acidic soils than in neutral or slightly alkaline soils. Strongly acid soils (pH four to five), though, can result in high concentrations of aluminum, iron and manganese in the soil solution, which may inhibit the growth of some plants.

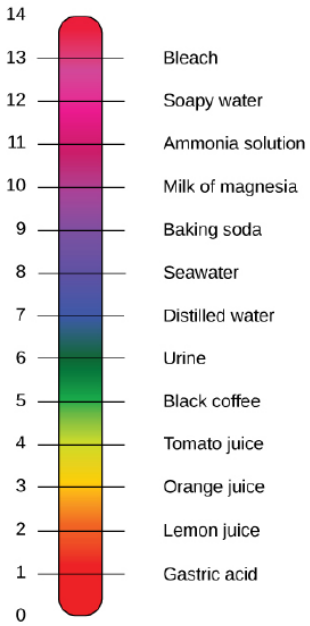


Figure 3.4.4.*b*: The pH scale measures the amount of hydrogen ions (H^+) in a substance. This shows the pH of common household solutions. Those with a pH greater than 7 are basic (alkaline), and those with a pH less than 7 are acidic. The solutions from high to low pH are bleach (13), soapy water (12), ammonia solution (11), milk of magnesia (10), baking soda (9), seawater (8), distilled water (7), urine (6), black coffee (5), tomato juice (4), orange juice (3), lemon juice (2), and gastric acid (1). Note that pH values are approximate. Image and caption (modified) from OpenStax/Edward Stevens (CC-BY). Access for free at openstax.org.

Several factors determine soil pH. Organic material in soil decreases pH to an extent, but it also acts as a **buffer**, limiting changes in pH. Climate is also important, with high amounts of rainfall increasing leaching and lowering pH. Some types of parent material, such as those high in silicon, decrease pH, while others, such as limestone increase pH.

Soil Composition

Soil consists of organic matter (about 5%), inorganic mineral matter (40-45% of soil volume), water (about 25%) and air (about 25%). The amount of each of the four major components of soil depends on the amount of vegetation, soil compaction, and water present in the soil.

The organic material consists of dead organisms in various stages of decomposition. It is dark-colored because it contains **humus**, partially decayed matter containing organic acids. Humus enriches the soil with nutrients, gives the soil a loose texture that holds

water, and allows air to diffuse through it. Oxygen is important for plant roots and many inhabitants of the soil. The organic component of soil serves as a cementing agent, returns nutrients to the plant, allows soil to store moisture, makes soil tillable for farming, and provides energy for soil microorganisms. Most soil microorganisms—bacteria, algae, or fungi—are dormant in dry soil, but become active once moisture is available.

The inorganic material of soil consists of rock, slowly broken down into smaller particles that vary in size. Soil particles that are 100 µm to 2 mm in diameter are **sand**. (A **micrometer**, µm, 10^{-6} m, or a millionth of a meter.) Soil particles between 2 and 100 µm are called **silt**, and even smaller particles, less than 2 µm in diameter, are called **clay**.

Soil should ideally contain 50 percent solid material and 50 percent pore space (figure 3.4.4.*c*). **Pore space** refers to the gaps in between soil particles. The larger the soil particles, the larger the pore spaces. Water can quickly pass through large pore spaces, so soils high in sand drain easily. Smaller soil particles have more surface area relative to volume and produce narrow pore spaces. Water clings to these surfaces, and soils high in clay thus retain water. (Clay is also negatively charged, which attracts water.) About one-half of the pore space should contain water, and the other half should contain air.

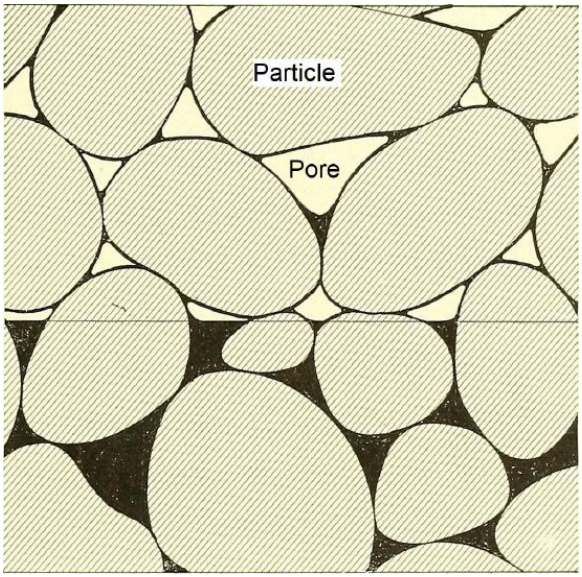


Figure 3.4.4.*c*: Soil particles and the pore spaces in between them. The black shading represents water, which fills the pore spaces in the lower portion of the soil and clings to narrow portions of pore space in the top portion of the soil. Image from "Forest physiography; physiography of the United States and principles of soils in relation to forestry" (public domain).

Soil texture is based on percentages of sand, silt, and clay (figure 3.4.4.*d*). Soils that have a high percentage of one particle size are named after that particle (a clay soil has a high percentage of clay). Other soils have a mixture of two particle sizes and very little of the third size. For example, silty clay has roughly 50% clay and 50% silt while sandy clay has 50-60% sand and 35-50% clay. Some soils have no dominant particle size and contain a mixture of sand, silt, and humus. These soils are called **loams**, and they are optimal for agriculture. A medium loam has roughly 40% sand, 40% silt, and 20% clay. Larger particles (sand) facilitate drainage, and small particles (clay) facilitate water retention, so loam soils both have good drainage and can remain moist. Soils that deviate slightly from a medium loam include loamy sand, sandy loam, sandy clay loam, clay loam, silty clay loam, and silty loam.

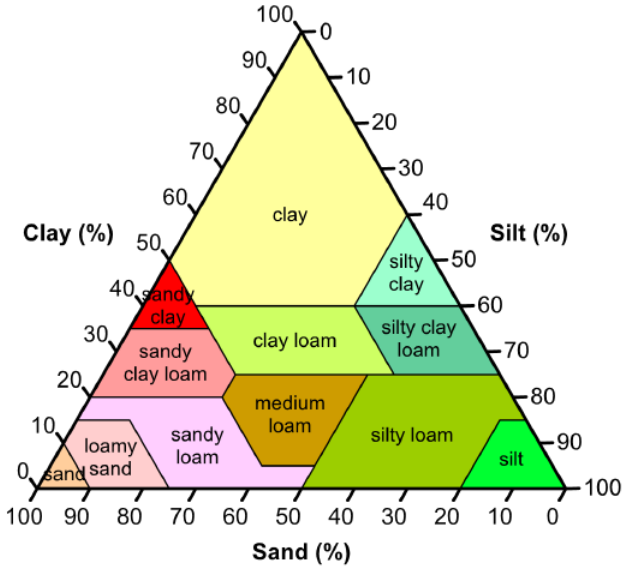


Figure 3.4.4.*d*: A soil texture triangle is used to determine the soil texture based on the percentages of sand, silt, and clay. The left axis represents the percentage of clay. The right axis represents the percentage of silt. The bottom axis represents the percentage of sand. Most soils that contain more than 40% are classified as clays. Exceptions are silty clays, which contain 40-60% clay and 40-60% silt. Another exception are sandy clays, which contain 35-50% clay and 50-70% sand. Clay loams and silty clay loams contain 25-40% clay. Clay loams contain 20-60% silt, and silty clay loams contain 60-75% silt. Sandy clay loams contain 20-35% clay and 50-80% sand. Medium loams contain 5-25% clay, 20-50% silt, and 30-55% sand. Soils with more than 50% silt and no more than 25% clay are either silty loams or silts. Silts have 80-100% silt and no more than 15% clay. Sand have 90-100% sand. Loamy sands have 75-90% sand and no more than 15% clay. A soil with a composition that does not fit into the already described categories would be a sandy loam. To determine the soil texture, first find the percentage of clay and trace a line horizontally to the right. At the same time, find the percentage of silt and trace a line diagonally (down and to the left, parallel to the clay axis). The two lines will meet at the correct soil type. For example, traced lines from 20% clay and 40% silt intersect at the medium loam category. To confirm, trace a diagonal line (down and to the right, parallel to the sit axis) from this point to the sand axis. This line cross the axis at 40%. In summary, a line can be traced from each axis that is parallel to axis that is counterclockwise of it. Image by Richard Wheeler/Zephyris at English Wikipedia (CC-BY-SA).

Organic Versus Mineral Soils

Soils can be divided into two groups based on how they form. **Organic soils** are those that are formed from sedimentation and often contain more than 30% organic matter. They form when organic matter, such as leaf litter, is deposited more quickly than it can be decomposed (figure 3.4.4.*e*). **Mineral soils** are formed from the weathering of rocks, typically contain no more than 30% organic matter, and are primarily composed of inorganic material. **Weathering** occurs when biological, physical, and chemical processes, such as erosion, leaching, or high temperatures, break down rocks.



Figure 3.4.4.*e*): Decomposition of organic material occurs very slowly in this acidic bog. The organic material accumulates, which is characteristic of organic soils. Image by William L. Farr (CC-BY-SA).

Soil Horizons

Soil distribution is not uniform because its formation results in the production of layers; together, the vertical section of a soil is called the **soil profile**. Within the soil profile, soil scientists define zones called horizons. A **horizon** is a soil layer with distinct physical and chemical properties that differ from those of other layers.

The soil profile has four distinct layers: 1) O horizon; 2) A horizon; 3) B horizon and 4) C horizon (figure 3.4.4.*f*-g). Upper horizons (labeled as the A and O horizons) are richer in organic material and so are important in plant growth, while deeper layers (such as the B and C horizons) retain more of the original features of the bedrock below. Some soils may have additional layers (like the E horizon, figure 3.4.4.*f*), or lack one of these layers. The thickness of the layers is also variable, and depends on the factors that influence soil formation. In general, immature soils may have O, A, and C horizons, whereas mature soils may display all of these, plus additional layers.

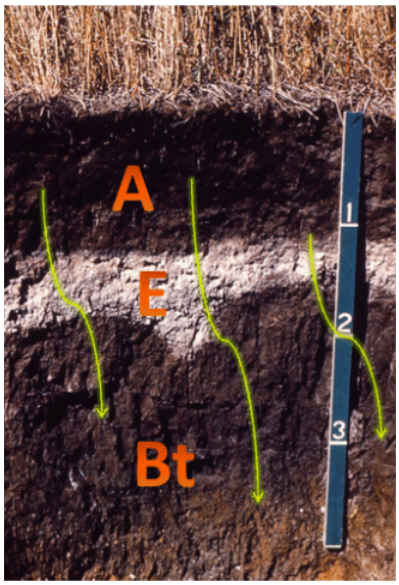


Figure 3.4.4.*f*): Soil Profile. Photograph shows a soil profile from South Dakota revealing three horizons. The A horizon (topsoil) is dark brown and extend about 1 foot. Below it is the white E (eluviated) horizon, which is about 6 inches thick. The B horizon (subsoil, labeled Bt) is the medium brown and the lowest visible layer The yellow arrows symbolize translocation of fine clays to the B horizon. The scale is in feet. Source: University of Idaho and modified by D. Grimley.

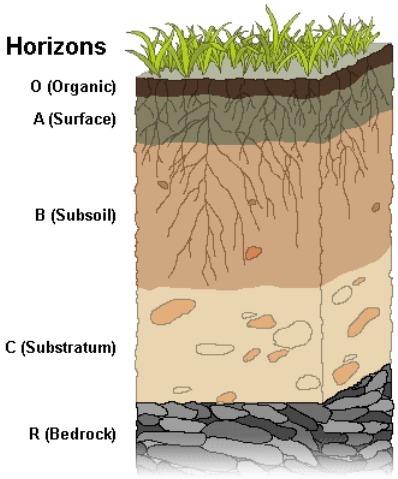


Figure 3.4.4.*g*): This image shows the different horizons, or layers, in soil. Vegetation grows at the surface. The top, deep brown layer is the O (organic) horizon. Below this is the gray A horizon (topsoil, or surface). Next is the light brown B horizon (subsoil). The C horizon (substratum, or soil base) is tan and includes larger pieces of rocks. The gray bedrock is the deepest layer. The E (eluviated) horizon that is sometimes found between the A and B horizons is absent here. This work by Wilsonbiggs is licensed under CC BY-SA 4.0

O horizon

The very top of the **O horizon** consists of partially decayed organic debris like leaves. This horizon is usually dark in color because of humus.

A horizon

The **A horizon** (topsoil) consists of a mixture of organic material with inorganic products of weathering, and it is therefore the beginning of true mineral soil. In this area, rainwater percolates through the soil and carries materials from the surface. The A horizon may be only 5 cm (2 in.), or it may over a meter. For instance, river deltas like the Mississippi River delta have deep layers of topsoil. Microbial processes occur in the top soil, and this horizon supports plant growth. Many organisms, such as earthworms and insects live among the plant roots in this horizon.

B horizon

The **B horizon** (subsoil) consists of small particles that have moved downward, resulting in a dense layer in the soil. In some soils, the B horizon contains nodules or a layer of calcium carbonate. The subsoil is usually lighter in color than topsoil and often contains an accumulation of minerals.

C horizon

The **C horizon** (soil base), includes the **parent material**, the organic and inorganic substances from which soils form. Weathering parent material represents the first steps in the chemical breakdown of rock into soil. Often the weathered parent material is underlain by the parent material itself, although in some places it has been carried from another location by wind, water, or glaciers. Beneath the C horizon lies bedrock. The chemical nature of the parent material, whether granite, limestone, or sandstone, for example, has a great influence on the fertility of the soil derived from it.

Factors Affecting Soil Formation and Composition

The fundamental factors that affect soil genesis can be categorized into five elements: climate, organisms, topography, parent material, and time. One could say that the relief, climate, and organisms dictate the local soil environment and act together to cause weathering and mixing of the soil parent material over time.

Climate

The role of climate in soil development includes aspects of temperature and precipitation. Soils in very cold areas with permafrost conditions (such as the arctic tundra) tend to be shallow and weakly developed due to the short growing season. In warm, tropical climates, soils tend to be thicker (but lacking in organic matter), with extensive leaching and mineral alteration. In such climates,

organic matter decomposition and chemical weathering occur at an accelerated rate. The presence of moisture and nutrients from weathering will also promote biological activity: a key component of a quality soil. See the Biomes chapter for more details about the effect of climate on soils.

Ancient soils, sometimes buried and preserved in the subsurface, are referred to as **paleosols** (figure 3.4.4.h) and reflect past climatic and environmental conditions.

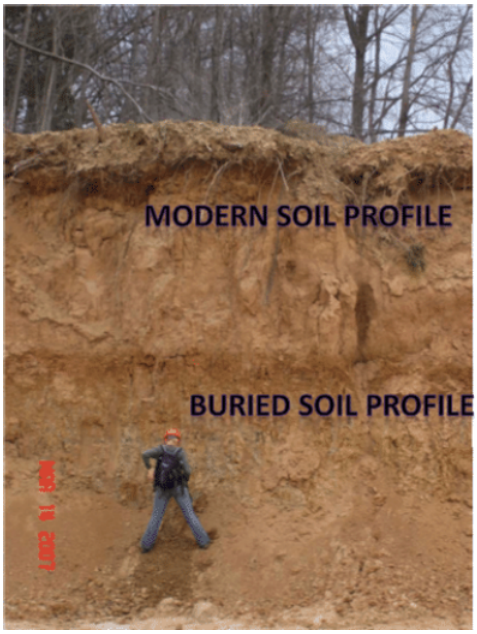


Figure 3.4.4.h: Modern versus Buried Soil Profiles. A buried soil profile, or paleosol (above geologist ‘s head), represents soil development during the last interglacial period. A modern soil profile (Alfisol) occurs near the land surface. Source: D. Grimley.

Organisms

The presence of living organisms in the soil (**soil biota**) greatly affects soil formation and structure. A diversity of animals are found in the soil such as nematodes, spiders, insects, centipedes, millipedes, pillbugs, slugs, and earthworms (figure 3.4.4.i). The soil also contains microorganisms like bacteria, archaea, fungi, and "protists". Animals and microorganisms can produce pores and crevices, and plant roots can penetrate into crevices to produce more fragmentation. Additionally, leaves and other material that fall from plants decompose and contribute to soil composition. Microorganisms not only decompose organic matter, but contribute to other processes in nutrient cycles, such as nitrogen fixation.

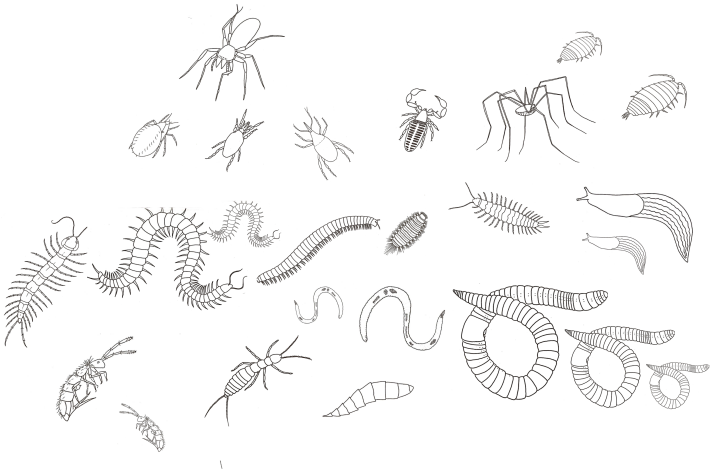


Figure 3.4.4.i: Illustration of animals commonly found in the soil. Image by Arne Hendriks (CC-BY)

Parent Material

Mineral soils form directly from the weathering of bedrock, the solid rock that lies beneath the soil, and therefore, they have a similar composition to the original rock. Other soils form in materials that came from elsewhere, such as sand and glacial drift. Materials located in the depth of the soil are relatively unchanged compared with the deposited material. Sediments in rivers may have different characteristics, depending on whether the stream moves quickly or slowly. A fast-moving river could have sediments of rocks and sand, whereas a slow-moving river could have fine-textured material, such as clay.

The type of parent material may also affect the rapidity of soil development. Parent materials that are highly weatherable (such as volcanic ash) will transform more quickly into highly developed soils, whereas parent materials that are quartz-rich, for example, will take longer to develop. Parent materials also provide nutrients to plants and can affect soil internal drainage.

Topography

Regional surface features (familiarly called “the lay of the land”) can have a major influence on the characteristics and fertility of a soil. Topography affects water runoff, which strips away parent material and affects plant growth. Soils on steep slopes are more prone to erosion and may be thinner than soils that are on relatively level ground. **Infiltration**, the percolating of water through the soil, is limited in steep soils.

The local topography can have important microclimatic effects. In the northern hemisphere, south-facing slopes are exposed to more direct sunlight angles and are thus warmer and drier than north-facing slopes. The cooler, moister north-facing slopes have a more dynamic plant community and thicker soils because extensive root systems stabilize the soil and reduce erosion (figure 3.4.4.j).



Figure 3.4.4.j: A north-facing slope (left) and south-facing slope (right) in the Mediterranean climate (chaparral) of the Santa Monica Mountains, California. Vegetation on the north-facing slope is denser due to cool, moist conditions relative to the south-facing slope. Image by Noah Elhardt (CC-BY).

Time

Time is an important factor in soil formation because soils develop over long periods. Soil formation is a dynamic process. Materials are deposited over time, decompose, and transform into other materials that can be used by living organisms or deposited onto the surface of the soil.

In general, soil profiles tend to become thicker (deeper), more developed, and more altered over time. However, the rate of change is greater for soils in youthful stages of development. The degree of soil alteration and deepening slows with time and at some point, after tens or hundreds of thousands of years, may approach an equilibrium condition where erosion and deepening (removals and additions) become balanced. **Young soils** (< 10,000 years old) are strongly influenced by parent material and typically develop horizons and character rapidly. Over time, as weathering processes deepen, mix, and alter the soil, the parent material becomes less recognizable as chemical, physical, and biological processes take effect. **Moderate age soils** (roughly 10,000 to 500,000 years old) are slowing in profile development and deepening, and may begin to approach equilibrium conditions. **Old soils** (>500,000 years

old) have generally reached their limit as far as soil horizonation and physical structure, but may continue to alter chemically or mineralogically.

Soil development is not always continual. Geologic events such as landslides, glacier advance, or the rising of shorelines can rapidly bury soils. Erosion in rivers and shorelines can cause removal or truncation of soils, and wind or flooding slowly deposit sediment that adds to the soil. Animals can mix the soil and sometimes cause **soil regression**, a reversal or "bump in the road" for the normal path of development, and this increases development over time.

Soil Taxonomy

Soils are classified into one of 12 soil orders based on soil horizons, how they form, and their chemical compositions. For example, Mollisols (figure 3.4.4.f), which are found in temperate grasslands, have a thick topsoil rich in organic content. Aridisols, on the other hand, are dry soils that contain calcium carbonate and are found in deserts. Each soil order is further divided into suborders. See USDA's The Twelve Orders of Soil Taxonomy and The Twelve Soil Orders from the University of Idaho for more details.

Attributions

Modified by Melissa Ha from the following sources:

- The Soil from *General Biology* by OpenStax (CC-BY)
- Soil Profiles and Processes and Soil-Plant Interactions from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Soil from *Biology* by John W. Kimball (licensed under CC-BY)

This page titled 3.4.4: Soils is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **7.4: Soils** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

3.4.5: Soil Degradation

Once fertile topsoil is lost, it is not easily replaced. **Soil degradation** refers to deterioration in the quality of the soil and the resultant reduction in its capacity to produce. Soils are degraded primarily by erosion, compaction, and salinization. Such processes often arise from poor soil management during agricultural activities. In extreme cases, soil degradation can lead to **desertification** (conversion of land to desert-like conditions) of croplands and rangelands in semi-arid regions. See the Sustainable Agriculture section for strategies to preserve soil quality (**soil conservation**).

Erosion

Erosion is the biggest cause of soil degradation. Soil productivity is reduced as a result of losses of nutrients, water holding capacity, and organic matter. **Water holding capacity** is a measure of of the soils ability to retain water. The two agents of erosion are wind and water, which act to remove the finer particles from the soil. **Wind erosion** occurs mostly in flat, dry areas and moist, sandy areas along bodies of water. Wind not only removes soil, but also dries and degrades the soil structure. **Water erosion** is the most prevalent type of erosion. It occurs when raindrops splash on the ground and when water moves down slope as a thin film, small streams, or a large stream.

Some amount of soil erosion is a natural process along sloping areas and/or in areas with soft or materials that do not stick together and are susceptible to movement by water, wind, or gravity. For instance, soil material can be mobilized in strong windstorms, along the banks of rivers, in landslides, or by wave action along coastlines. However, human activities such as construction, logging, and off-road vehicle use promote erosion by removing the natural vegetation cover protecting the soil. Agricultural practices such as overgrazing and leaving plowed fields bare for extended periods contribute to farmland erosion. Each year, an estimated two billion metric tons of soil are eroded from farmlands in the United States alone. The soil transported by the erosion processes can also create problems elsewhere (for example, by clogging waterways and filling ditches and low-lying land areas). The areas most vulnerable to soil erosion include locations with thin organic (A and O) horizons and hilly terrains (figure $3.4.5.a$).

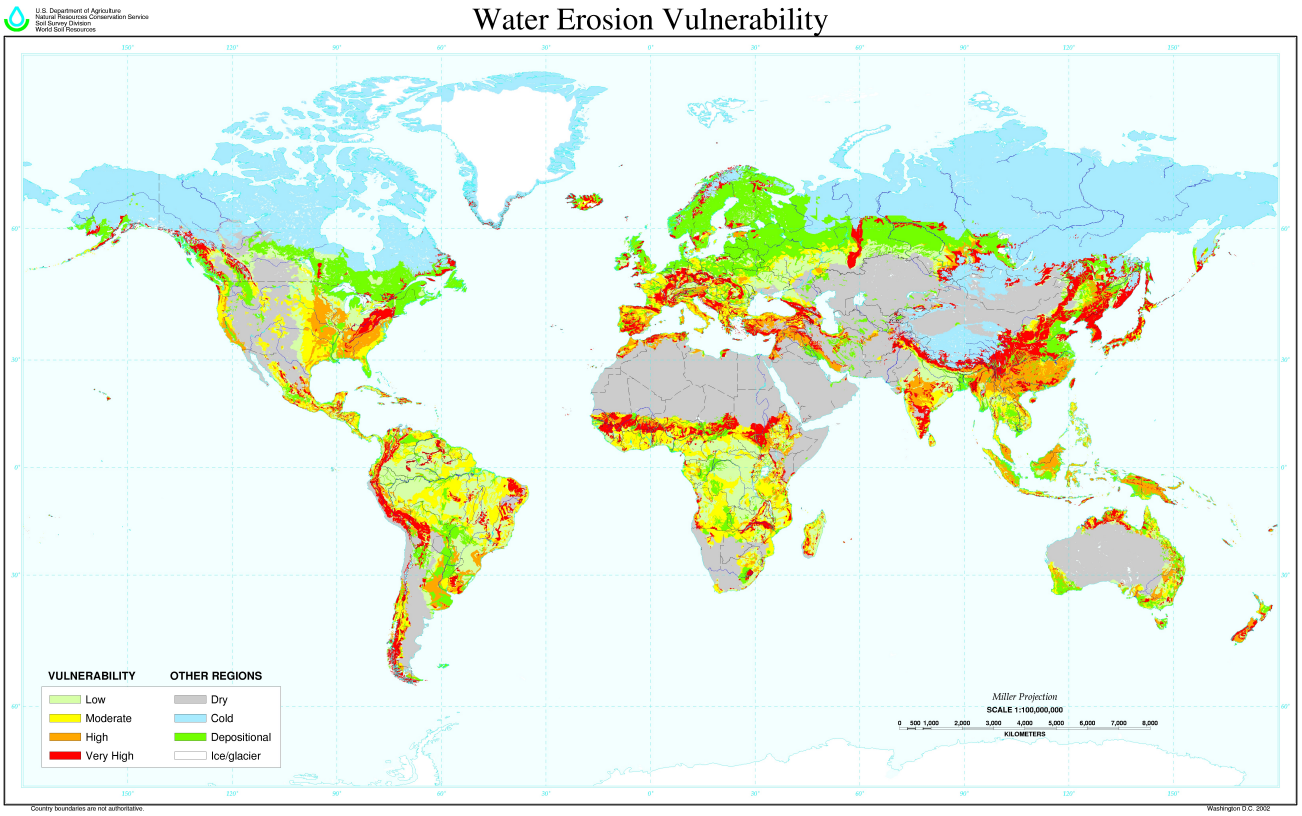


Figure 3.4.5.*a*: Global map of water erosion vulnerability. There are four levels of vulnerability: very high (red), high (orange), moderate (yellow), and low (light green). Examples of regions with very high vulnerability to erosion by water include the west coast of South America, parts of the eastern United States, regions around the Mediterranean Sea, a strip across western and central Africa, Korean and Japan, Northern Australia, and New Zealand. Examples of regions with low vulnerability include northern South America, the Democratic Republic of Congo, northern India and Nepal, and western Russia. Other regions are labeled as dry (gray), cold (light blue), depositional (bright green), or ice/glacier (white). For example, northern regions are cold and ice/glacier. Much of the western United States, Chile and Argentina, northern Africa, the Middle East, Kazakhstan, Mongolia, and much of Australia are considered dry. The largest depositional area (region where sediments accumulate) is northeastern Europe. Image by USDA-NRCS, Soil Science Division, World Soil Resources (public domain).

Compaction

In modern agricultural practices, heavy machinery is used to prepare the seedbed, for planting, to control weeds, and to harvest the crop. The use of heavy equipment has many advantages in saving time and labor, but can cause compaction of soil and disruption of the natural soil biota. Much compaction is reversible and some is unavoidable with modern practices; however, serious compaction issues can occur when the equipment is used excessively during times when the soil has a high water content. The problem with soil compaction is that increased soil density limits root penetration depth and may inhibit proper plant growth.

Salinization

When considerable quantities of salt accumulate in the soil in a process known as **salinization**, many plants are unable to grow properly or even survive. This is especially a problem in irrigated farmland. Groundwater used for irrigation contains small amounts of dissolved salts. Irrigation water that is not absorbed into the soil evaporates, leaving the salts behind. This process repeats itself and eventually severe salinization of the soil occurs. A related problem is water logging of the soil. When cropland is irrigated with excessive amounts of water in order to leach salts that have accumulated in the soil, the excess water is sometimes unable to drain away properly. In this case it accumulates underground and causes a rise in the subsurface water table. If the saline water rises to the level of the plant roots, plant growth is inhibited.

Desertification

Land that was previously suited for growing crops may be turned into desert by climate change and the activities of humans, such as poor farming practices, livestock overgrazing, and overuse of available water. This process, called **desertification**, is a serious problem worldwide. Plants and soil types that are non-arid (not dry) specifically help water soak into the ground (infiltration) and water retention. When desertification begins, it leads to a reduction in vegetation and degraded soil quality, and this further increases the aridity and spreads the desert via a **positive feedback loop** (meaning that the processes feed on themselves promoting an increasing spiral).

Figure 3.4.5.*b* shows areas of the world and their vulnerability to desertification. Note the red and orange areas in the western and midwestern United States. The Dust Bowl of the 1930s is a classic example of human-caused desertification (figure 3.4.5.*c*). Poor cultivation and grazing practices -- coupled with severe drought conditions -- led to severe wind erosion of soil in a region of the Great Plains that became known as the "Dust Bowl". Wind stripped large areas of farmlands of topsoil, and formed clouds of dust that traveled as far as the eastern United States.

Sometimes there is a conflict between what is known to prevent desertification and what an individual farmer feels is needed to do to make a living. Mitigating the desertification process includes both societal steps and individual education on alternatives.

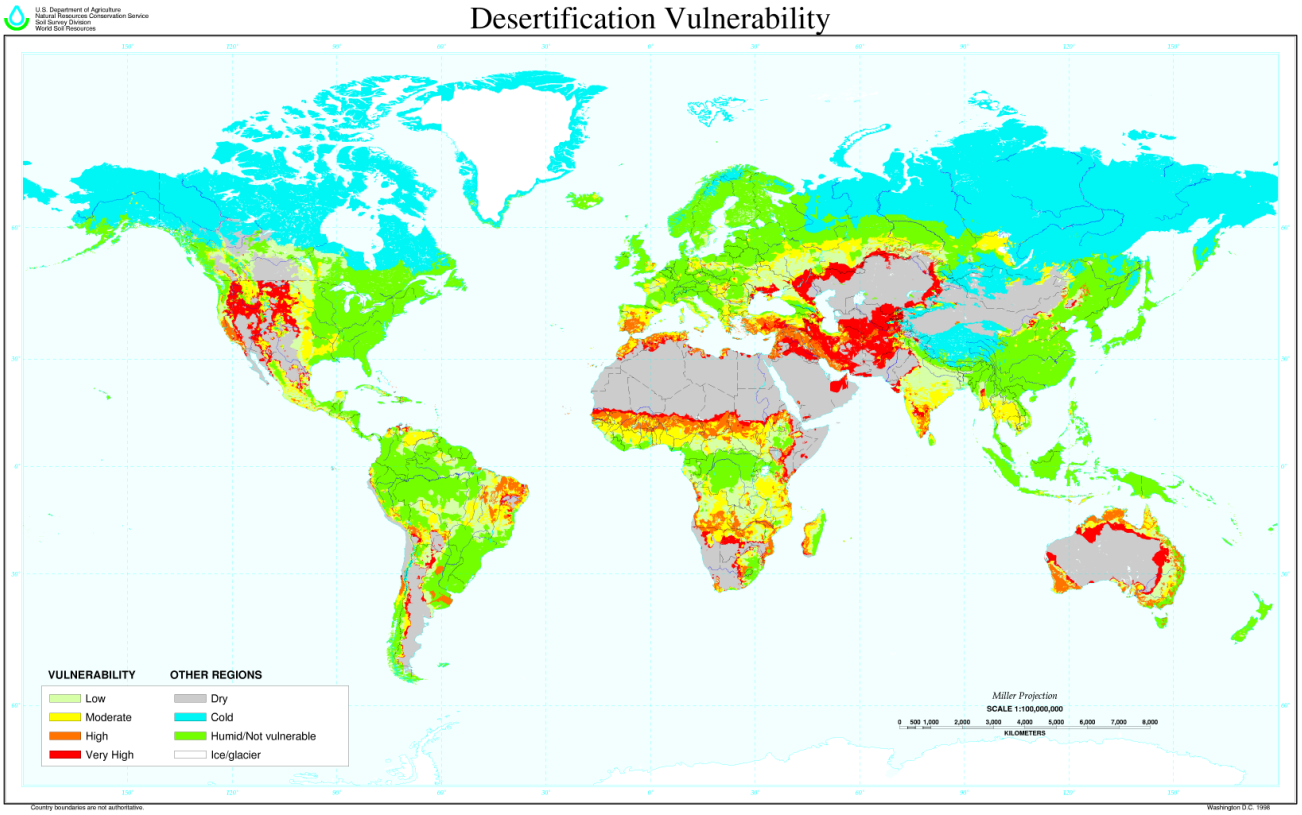


Figure 3.4.5.*b*: World map showing desertification vulnerability. There are four levels of vulnerability: very high (red), high (orange), moderate (yellow), and low (light green). Examples of regions with a very high vulnerability to desertification include the northwestern United States, a strip through western and central Africa, parts of the middle East, and parts of Australia. Examples of regions with low vulnerability include parts of Canada, parts of western Europe, and parts of Angola and the Democratic Republic of Congo. Other regions are labeled as dry (gray), cold (light blue), humid/not vulnerable (bright green), or ice/glacier (white). For example, northern regions are cold and ice/glacier. Some of southwestern United States, Chile and Argentina, northern Africa, the Middle East, Kazakhstan, Mongolia, and much of Australia are considered dry. Examples of humid/not vulnerable areas include the eastern United States, northern and eastern South America, western Europe, Central Africa, and East Asia. Image by USDA (public domain).



Figure 3.4.5.*c*: A cloud of dust in Kansas in 1935, part of the Dust Bowl. Image by NOAA Photo Library (public domain).

Attribution

Modified by Melissa Ha from the following sources:

- The Great Basin and the Basin and the Range from *An Introduction to Geology* by Chris Johnson et al. (licensed under CC BY-NC-SA)
- Soils from *AP Environmental Science* by University of California College Prep, University of California (licensed under CC-BY). Download for free at CNX.
- Soil and Sustainability from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors (licensed under CC-BY). Download for free at CNX.

This page titled 3.4.5: Soil Degradation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **7.5: Soil Degradation** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

3.4.6: Data Dive- Beaver Impacts on Wetlands

Overview

An ecosystem engineer is any animal that creates, significantly modifies, maintains or destroys a habitat. These organisms impact both the abiotic and biotic components of a habitat and can thus completely change landscapes. One 2017 study sought to understand if beavers (classic ecosystem engineer) could even be used as a tool for habitat restoration of wetlands. Specifically, this study focused on quantifying if wetland biodiversity improved as a result of reintroducing beavers on a landscape degraded by agriculture. Figure 2.4.6a below displays some of the results in this study:

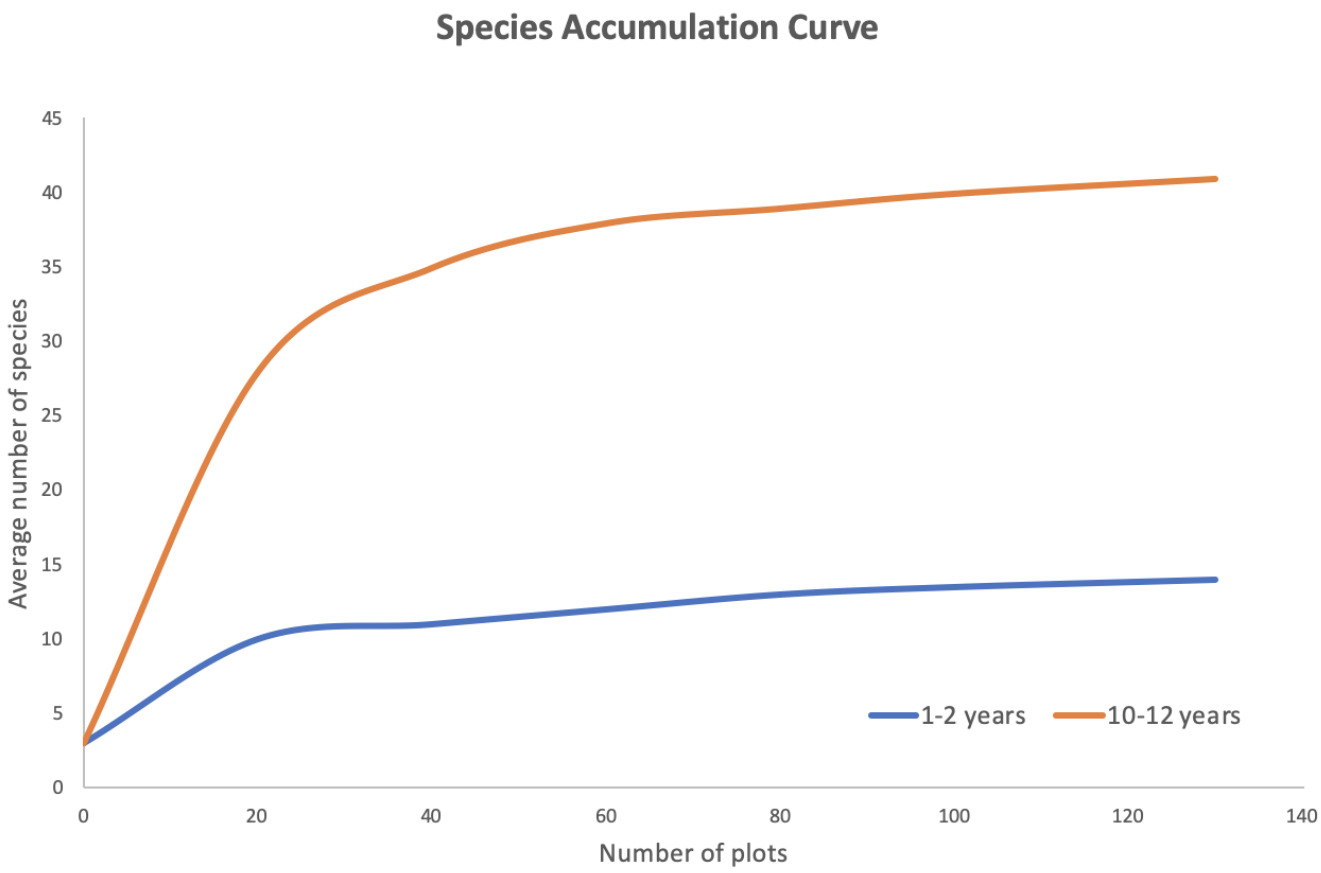


Figure 3.4.6.a: Average number of species observed in sample plots 1-2 years and 10-12 years after beavers were introduced. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Law A, Graywood MJ, Jones KC, Ramsay P, and Willby NJ 2017.

Questions

1. What is the independent (explanatory) variable and the dependent (response) variable?
2. What question(s) are the authors trying to answer with this graph?
3. What trend(s) can be observed in this graph between the 1-2 and 10-12 timetables? Support your answer by referring to appropriate patterns in the graph.
4. Do you think like the authors are satisfied with the results in the graph? Why?
5. How can the results of this graph to inform future reintroduction of beavers where wetland restoration is needed?
6. What information/patterns is not clear from this graph?

Raw Data From Above Graph(s)

Table 3.4.6.*a***:** Raw data for average number of species observed in sample plots 1-2 years and 10-12 years after beavers were introduced. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Law A, Graywood MJ, Jones KC, Ramsay P, and Willby NJ 2017.

Number Of Plots	Average Number Of Species After 1-2 Years	Average Number Of Species After 10-12 Years
0	3	3
20	10	28
40	11	35
60	12	38
80	13	39
100	13.5	40
130	14	41

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 3.4.6: Data Dive- Beaver Impacts on Wetlands is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **7.6: Data Dive- Beaver Impacts on Wetlands** by Melissa Ha and Rachel Schleiger has no license indicated.

3.4.7: Review

Summary

After completing this chapter you should be able to...

- Differentiate between abiotic and biotic ecosystem components.
- Describe the three main categories of ecosystems.
- Explain the law of conservation of mass.
- Discuss the biogeochemical cycles of carbon, nitrogen, phosphorus, and sulfur.
- Explain how human activities have impacted these cycles and the potential consequences for Earth.
- Explain how soil characteristics influence plant growth.
- Identify and describe each component of soil.
- Distinguish among sand, silt, and clay and explain how particle size influences soil texture.
- Describe each horizon in a typical soil profile.
- Explain how soils are formed, describing each of the five major factors that affect soil formation and composition.
- Describe the major types and causes of soil degredation.

Ecosystems consist of living (**biotic**) and nonliving (**abiotic**) components. They can be classified as freshwater, marine, or terrestrial. **Resistance** and **resilience** are measures of ecosystem health.

Matter is anything that occupies space and has mass. Pure forms of matter are called **elements** and the smallest units of an element are **atoms**. Atoms form molecules through **ionic, covalent,** or **hydrogen bonding**. Molecules that contain carbon and hydrogen covalent bonds are called **organic**. There are four main type of large organic molecules (**biological macromolecules**) in organisms: carbohydrates, lipids, proteins, and nucleic acids.

The chemical elements that organisms need continuously cycle through ecosystems. Cycles of matter are called **biogeochemical cycles,** or **nutrient cycles**, because they include both biotic and abiotic components and processes. Examples of biogeochemical cycles include the carbon, nitrogen, phosphorus, and sulfur cycles, and each of these can be altered through human activities.

Soil consists of organic and inorganic material as well as water and air. The organic material of soil is made of **humus**, which improves soil structure and provides nutrients. Soil inorganic material consists of rock slowly broken down into smaller particles that vary in size, such as sand, silt, and loam. Soils form slowly as a result of biological, physical, and chemical processes. Soil is not homogenous because its formation results in the production of layers called a **soil profile**. Most soils have four distinct **horizons,** or layers: O, A, B, and C. Their composition is influenced by the climate, presence of living organisms, topography, parent material, and time. The processes of erosion, compaction, and desertification degrade soils. While these processes occur naturally to an extent, they are exacerbated by certain agricultural practices, deforestation, and other human activities.

Attribution

Modified by Melissa Ha from the following sources:

- The Soil and Biogeochemical Cycles from *General Biology* by OpenStax (licensed under CC-BY)
- Nutrient Cycles from *Human Biology* by Suzanne Wakim and Mandeep Grewal (CC-BY-NC)

This page titled 3.4.7: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **7.7: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

3.5: Biomes

Chapter Hook

At the North Pole there is nothing but ice and snow. The taiga (also known as the boreal forest) is the northernmost place that trees can grow. At first, trees are very sparse and stunted. Sometimes these trees can take upwards of 50 years to get bigger than a seedling. Eventually, the taiga turns into a sea of trees. There are as many trees here as in all rainforests combined, representing nearly a third of all trees on Earth. As such, the taiga is the largest carbon sink on land and contributes significantly to the contribution of oxygen in the atmosphere. The taiga is just one of Earths biomes, each is unique and extraordinary in their own ways.

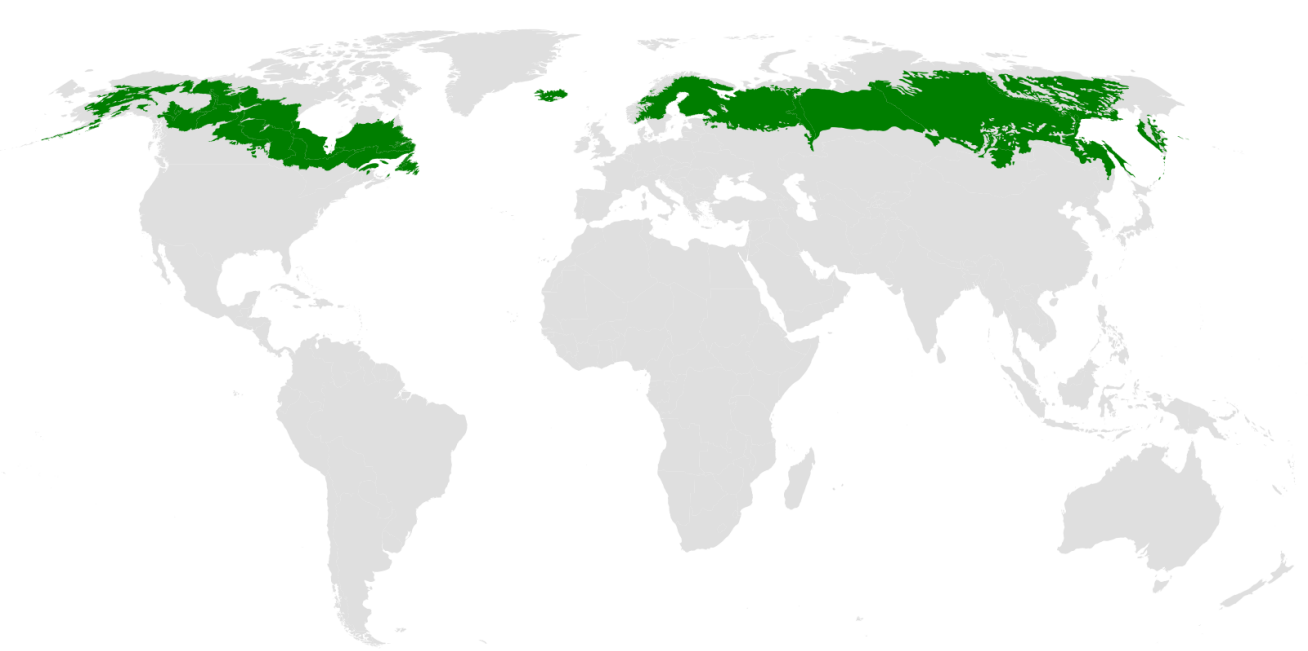


Figure 3.5.*a* World map with shading indicating the location of the taiga biome. Image by Mark Baldwin-Smith (licensed under CC-BY-SA 4.0)

Attribution

Rachel Schleiger (CC-BY-NC)

3.5.1: Climate and Biomes

3.5.2: Terrestrial Biomes

3.5.3: Aquatic Biomes

3.5.4: Data Dive- Biome Carbon Storage

3.5.5: Review

This page titled 3.5: Biomes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

3.5.1: Climate and Biomes

Biomes are large-scale environments that are distinguished by characteristic climate and vegetation (figure 3.5.1.a). Biomes are also characterized by the animals and other organisms there, which are influenced by vegetation and climate patterns. The Earth's biomes are categorized into two major groups: terrestrial and aquatic. **Terrestrial biomes** are based on land, while **aquatic biomes** include both ocean and freshwater biomes. Altitude and latitude, which affect temperature and precipitation determine the distribution of biomes.

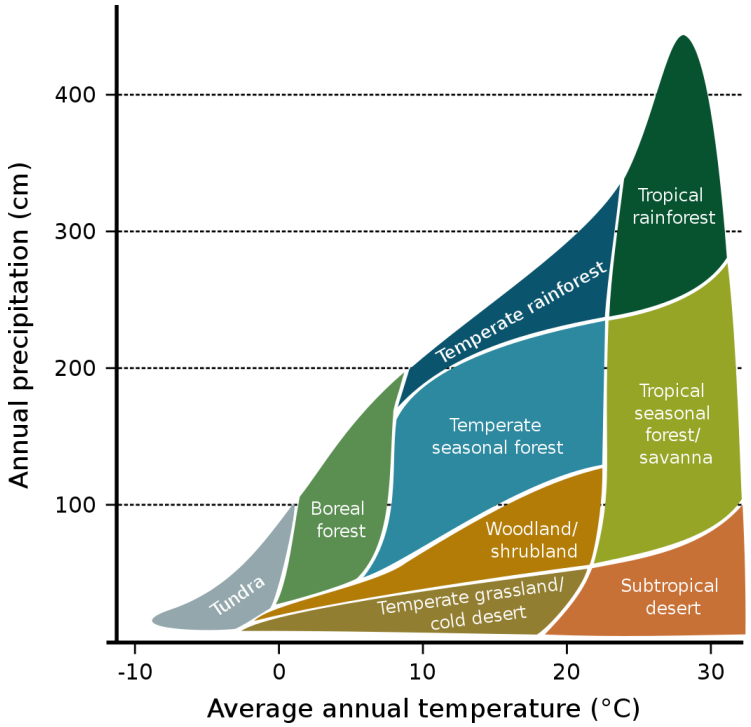


Figure 3.5.1.a: Average annual temperature and annual precipitation are two climatic factors that determine the distribution of biomes. This graph shows annual precipitation in centimeters on the y-axis and average annual temperature in degrees Celsius on the x-axis. Tropical rainforests have the highest temperatures and most precipitation. Tropical seasonal forests and savannas have high temperatures (20-30°C) with precipitation from approximately 50 to 270 centimeters. Subtropical deserts have similarly high temperatures but dry conditions. Temperate grasslands and cold deserts range in average annual temperature from below 0 to 22°C and have less than 50 centimeters of precipitation per year. Woodlands/shrublands (such as the chaparral) have the same temperature range as cold deserts but can receive a little over 100 cm of precipitation per year. Temperate forests have intermediate temperatures (approximately 16-22°C), but temperate rainforests receive more precipitation than temperate seasonal forests. Boreal forests receive a little less precipitation than temperate forests, but are colder (approximately 1-10°C). The tundra has the coldest temperatures and little precipitation. Deserts have a range of temperatures but low precipitation. Image by Navarras (public domain).

Low latitudes (near the equator) have high temperatures, while high latitudes (near the poles) have low temperatures. This is because the sun hits the equator more directly. Sunlight hits the poles at an angle, reducing the intensity of light (and heat energy) per unit of area. Temperature also decreases with altitude. At high altitudes, the atmosphere is thinner and traps less heat energy from the sun. Because temperatures decline with altitude as well as latitude, similar biomes exist on mountains even when they are at low latitudes. As a rule of thumb, a climb of 1000 feet (about 300 m) is equivalent in changed flora and fauna to a trip northward of some 600 miles (966 km).

Where precipitation is moderately abundant — 40 inches (about 1 m) or more per year — and distributed fairly evenly throughout the year, the major determinant is temperature. It is not simply a matter of average temperature, but includes such limiting factors as whether it ever freezes or length of the growing season. Biomes are thus characterized not only by average temperature and precipitation but also their seasonality.

Not only does latitude influence temperature, but it also affects precipitation. For example, deserts tend to occur at latitudes of around 30° and at the poles, both north and south, driven by circulation and prevailing wind patterns in the atmosphere. The engine

that drives circulation in the atmosphere and oceans is solar energy, which is determined by the average position of the sun over the Earth's surface. Direct light provides uneven heating depending on latitude and angle of incidence, with high solar energy in the tropics, and little or no energy at the poles. Atmospheric circulation and geographic location are the primary causal agents of deserts. At approximately 30° north and south of the equator, sinking air produces **trade wind deserts** like the Sahara and the Outback of Australia (figure 3.5.1.*b*).

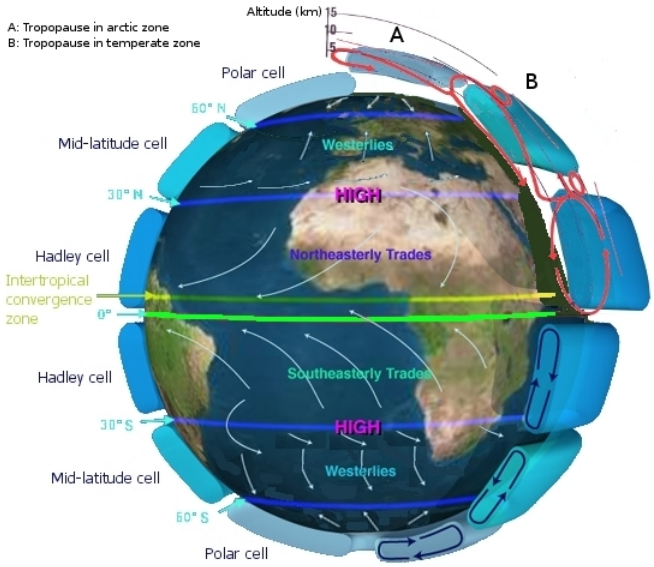


Figure 3.5.1.*b*: The generalized circulation of the atmosphere. Solar energy falling on the equatorial belt heats the air and causes it to rise. The rising air cools and its contained moisture falls back on the tropics as rain. The drier air then continues to spread toward the north and south where it sinks back at about 30 degrees north and south latitudes. This sinking drier air creates belts of predominant high pressure along which desert conditions prevail. These belts of predominantly high pressure have air that descends along these belts and flows either north to become the westerlies or south to become the trade winds. Note the arrows indicating general directions of winds in the latitude zones. The trade winds are predominant in the tropics and the westerlies in the mid-latitudes.

The MinuteEarth video below discusses the global climate patterns which lead to deserts.



Rainshadow deserts are produced where prevailing winds with moist air dries as it is forced to rise over mountains. The prevailing winds in the western half of North America blow in from the Pacific laden with moisture. Each time this air rises up from the western slopes of, successively, the Coast Ranges, the Sierras and Cascades, and finally the Rockies, it cools and its capacity to hold moisture decreases. The excess moisture condenses to rain or snow, which drenches the mountain slopes beneath. When the air reaches the eastern slopes, it is relatively dry, and much less precipitation falls. This phenomenon is called the **rainshadow effect** (figure 3.5.1.*c*). How much rain falls and when influences the type of biome. For example, the **Great Basin Desert** (figure 3.5.1.*d*) is a rain shadow desert produced as moist air from the Pacific rises by lifting over the Sierra Nevada Mountain (and other) and loses moisture from previous condensation and precipitation on the rainy side of the range(s).

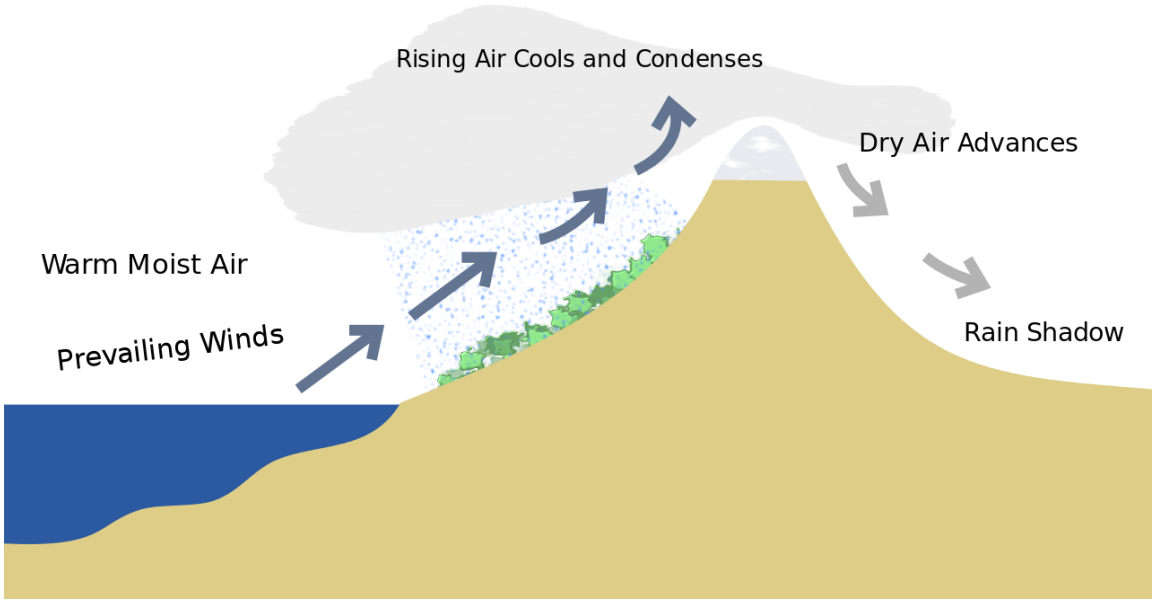


Figure 3.5.1.*c*: The rain shadow effect. Warm, moist air is carried up a mountain by prevailing winds. The rising air cools and condenses, resulting in precipitation as it moves up the mountain. The dry air advances down the other side of the mountain, resulting in a rain shadow (a dry region). Image by domdomegg (CC-BY).



Figure 3.5.1.*d*: Map of the Great Basin Desert. Image by USGS (public domain).

Attributions

Modified by Melissa Ha from the following sources:

- Biomes from *Biology* by John W. Kimball (licensed under CC-BY)
- Prelude to Deserts and The Origin of Deserts from *An Introduction to Geology* by Chris Johnson et al. (licensed under CC BY-NC-SA)
- Community Ecology and Terrestrial Biomes from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Terrestrial Biomes from *Biology 2e* by OpenStax (CC-BY). Access for free at openstax.org.

This page titled 3.5.1: Climate and Biomes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **8.1: Climate and Biomes** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

3.5.2: Terrestrial Biomes

Because each biome is defined by climate, the same biome can occur in geographically distinct areas with similar climates (figure 3.5.2.a). You can further explore the distribution of biomes using the HHMI Biome Viewer.

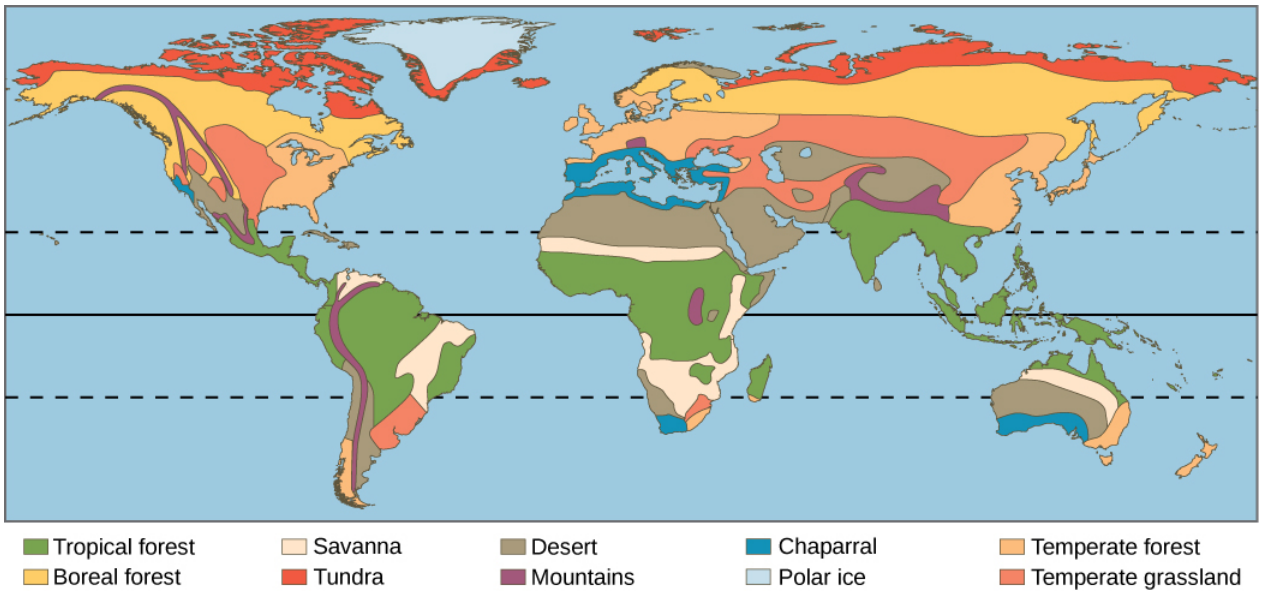


Figure 3.5.2.a: Each of the world's eight major biomes is distinguished by characteristic temperatures and amount of precipitation. Tropical forests are found closest to the equator (solid black line). Savannas are just north or south of tropical forests. Deserts occur at 30 degrees north and south (dashed black line). Chaparral is in southern California, around the Mediterranean sea, at the southern tip of Africa, and along the southwestern Australian coast. Temperate forests are found the eastern United States, tip of South America, western Europe, northeastern Asia, eastern Australia, and New Zealand. Boreal forests are found in Canada, Scandinavian countries, and Russia. Temperate grasslands are found in western South America, the central United States, western Europe, and central Asia. Tundra is at the north of the map. Much of Greenland is polar ice. Key mountains include the Sierra Nevada, Rocky Mountains, Andes, and Mitumba Mountains.

There are eight major terrestrial biomes: tropical rainforests, savannas, deserts, chaparral, temperate grasslands, temperate forests, taiga (boreal forests), and Arctic tundra.

Tropical Rainforests

Tropical rainforests, found in equatorial regions (figure 3.5.2.a), are the most biodiverse terrestrial biome (figure 3.5.2.b). In the Western Hemisphere, the tropical rain forest reaches its fullest development in the jungles of Central and South America. The trees are very tall and of a great variety of species. One rarely finds two trees of the same species growing close to one another. Most of the animals — mammals and reptiles, as well as birds and insects — live in the trees.



Figure 3.5.2.*b*: Species diversity is very high in tropical wet forests, such as these forests of Madre de Dios, Peru, near the Amazon River. (credit: Roosevelt Garcia)

This biodiversity is under extraordinary threat primarily through logging and deforestation for agriculture. Tropical rainforests have also been described as nature's pharmacy because of the potential for new drugs that is largely hidden in the chemicals produced by the huge diversity of plants, animals, and other organisms. The vegetation is characterized by plants with spreading roots and broad leaves. The broad-leafed trees are mostly **evergreen**, with leaves that persist throughout the year, but in tropical seasonal forests (see below), some are **deciduous,** with leaves falling off in the dry season.

The temperature and sunlight profiles of tropical rainforests are stable in comparison to other terrestrial biomes, with average temperatures ranging from 20°C to 34°C (68°F to 93°F). Month-to-month temperatures are relatively constant in tropical rainforests, in contrast to forests farther from the equator. This lack of temperature seasonality leads to year-round plant growth rather than just seasonal growth. In contrast to other ecosystems, a consistent daily amount of sunlight (11–12 hours per day year-round) provides more solar radiation and therefore more opportunity for **net primary productivity** (a measure of how carbon is stored by the photosynthesis that occurs in a community).

The annual rainfall in tropical rainforests ranges from 125 to 660 cm (50–200 in). Some tropical rainforests have relatively consistent rainfall whereas others (the **tropical seasonal forests**) have distinct wet months, in which there can be more than 30 cm (11–12 in) of precipitation, as well as dry months in which there are fewer than 10 cm (3.5 in) of rainfall. Nevertheless, the driest month of a tropical rainforest can still exceed the annual rainfall of some other biomes, such as deserts.

Tropical rainforests are characterized by vertical layering of vegetation and the formation of distinct habitats for animals within each layer. The vegetation is so dense that little light reaches the forest floor. On the forest floor is a sparse layer of plants and decaying plant matter. Above that is an **understory** of short, shrubby foliage. A layer of trees rises above this understory and is topped by a closed upper canopy—the uppermost overhead layer of branches and leaves. Some additional trees emerge through this closed upper canopy. These layers provide diverse and complex habitats for the variety of plants, animals, and other organisms. Many species of animals use the variety of plants and the complex structure of the tropical wet forests for food and shelter. Some organisms live several meters above ground, rarely descending to the forest floor.

Epiphytes

Epiphytes are plants that live perched on sturdier plants. They do not take nourishment from their host as parasitic plants do. Because their roots do not reach the ground, they depend on the air to bring them moisture and inorganic nutrients. Examples of epiphytes include many orchids, ferns (figure 3.5.2.*c*), and bromeliads (members of the pineapple family like "Spanish moss").



Figure 3.5.2.c: Epiphytes, graceful fern (*Niphidium crassifolium*), grow at the Soberania National Park in Panama. Image by Katja Schulz (CC-BY)

Tropical Soils

While productivity is high in tropical rainforests, the soils themselves tend to be of very poor quality. Because of the high rainfall, nutrients are quickly washed out of the topsoil unless they are incorporated in the forest plants. As plant and animal debris falls to the ground, it is quickly decomposed because of the warmth and moisture there. Thus minerals are found mainly in the forest plants, not in the soil. When the plants are removed and cultivation attempted, the soils quickly lose fertility. The situation is made worse by the lack of humus. Additionally, the topsoil may be no thicker than 5 cm (~2 inches), and most of these soils have high iron and aluminum content. Once exposed to the sun, these soils quickly bake into a bricklike material that cannot be cultivated.

Savannas

Savannas have mostly **perennial, herbaceous** plants (non-woody plants that persist for more than one year) with scattered trees, and they are located in Africa, South America, and northern Australia. Savannas are usually hot, tropical areas with temperatures averaging from 24 °C to 29 °C (75 °F to 84 °F) and an annual rainfall of 10–40 cm (3.9–15.7 in). Savannas have an extensive dry season; for this reason, forest trees do not grow as well as they do in the tropical rainforest. As a result, grasses and **forbs** (herbaceous plants other than grasses) are dominant (figure 3.5.2.$d-e$). Since fire is an important source of disturbance in this biome, plants have evolved well-developed root systems that allow them to quickly resprout after a fire. Depending on the region, large grazing animals like giraffes, zebras, and gazelles and their predators (cheetahs, lions, etc.) are found in savannas.



Figure 3.5.2.*d*: Savannas, like this one in Taita Hills Wildlife Sanctuary in Kenya, are dominated by grasses. (credit: Christopher T. Cooper)



Figure 3.5.2.*e*: Although savannas are dominated by grasses, small woodlands, such as this one in Mount Archer National Park in Queensland, Australia, may dot the landscape. (credit: "Ethel Aardvark"/Wikimedia Commons)

Deserts

Deserts exist between 15^{o} and 30^{o} north and south latitude. As discussed in the previous section (Climate and Biomes), deserts are frequently located on the downwind, or lee side, of mountain ranges due to the rain shadow effect. This is typical of the North American deserts, such as the Mohave and Sonoran deserts and the previously mentioned Great Basin Desert. Deserts in other regions (tradewind deserts), such as the Sahara and the Outback of Australia (both previously mentioned) as well as the Namib Desert in southwestern Africa and are dry because of the high-pressure, dry air descending at those latitudes.

Deserts are very dry; evaporation typically exceeds precipitation. Deserts are characterized by low annual precipitation of fewer than 30 cm (12 in) with little monthly variation and lack of predictability in rainfall. Some years may receive tiny amounts of rainfall, while others receive more. In some cases, the annual rainfall can be as low as 2 cm (0.8 in) in deserts located in central Australia ("the Outback") and northern Africa.

Hot deserts (subtropical deserts) can have daytime soil surface temperatures above 60°C (140°F) and nighttime temperatures approaching 0°C (32°F). **Cold deserts** that experience freezing temperatures during the winter and any precipitation is in the form of snowfall. The largest of these deserts are the Gobi Desert in northern China and southern Mongolia, the Taklimakan Desert in western China, the Turkestan Desert, and the Great Basin Desert of the United States.

The low species diversity of this biome is closely related to its low and unpredictable precipitation. Despite the relatively low diversity, desert species exhibit fascinating adaptations to the harshness of their environment. Very dry deserts lack perennial vegetation that lives from one year to the next; instead, many plants are annuals that grow quickly and reproduce when rainfall does occur, then they die. Perennial plants in deserts are characterized by adaptations that conserve water: deep roots, reduced foliage, and water-storing stems (figure 3.5.2.*f*). Seed plants in the desert produce seeds that can lie dormant for extended periods between rains. Most animal life in deserts has adapted to a nocturnal life, spending the hot daytime hours beneath the ground. The Namib Desert is the oldest on the planet, and has probably been dry for more than 55 million years. It supports a number of endemic species (species found only there) because of this great age. For example, the unusual gymnosperm *Welwitschia mirabilis*is the only extant species of an entire order of plants. There are also five species of reptiles considered endemic to the Namib.



Figure 3.5.2.*f*: Many desert plants have tiny leaves or no leaves at all to reduce water loss. The leaves of ocotillo, shown here in the Chihuahuan Desert in Big Bend National Park, Texas, appear only after rainfall and then are shed. (credit "bare ocotillo": "Leaflet"/Wikimedia Commons)

Many of the animals in the desert (mammals, lizards and snakes, insects, and even some birds) are adapted for burrowing to escape the scorching heat of the desert sun. Many of them limit their forays for food to the night.

The net primary productivity of the desert is low. High productivity can sometimes be achieved with **irrigation**, but these gains are often only temporary because it leads to the buildup of **salts** and **minerals**. Large areas of formerly unproductive desert in the United States, Israel, and Egypt have been converted into fertile fields through irrigation. However, even the best irrigation water contains dissolved salts. Because the rainfall in deserts is so low, any water that does not immediately run off remains near the surface and is largely lost by evaporation. The salts it carries are left near the top of the soil. Their accumulation may make the soil so alkaline (basic) and so salty that it is prohibitive to agriculture (figure 3.5.2.*g*). In the U. S., the situation is especially severe in the Great Basin because water flowing down from the mountains—bearing its load of dissolved salts—cannot flow on to the ocean but simply flows out onto the valley floors and evaporates.



Figure 3.5.2.*g*: Salt formations at the Badwater Basin in Death Valley, California. Minerals accumulate when water pools in this low-lying area, making conditions impossible for agriculture. Some plants, like pickleweed and saltgrass, are adapted to grow around the edges of the salt flats. Image by "Mike" Michael L. Baird (CC-BY).

Mediterranean Scrub (Chaparral)

The **Mediterranean scrub**, is also called chaparral or scrub forest and is found in California, along the Mediterranean Sea, and along the southern coast of Australia (figure 3.5.2.*h*). The annual rainfall in this biome ranges from 65 cm to 75 cm (25.6–29.5 in) and the majority of the rain falls in the winter. Summers are very dry and many chaparral plants are dormant during the summertime.



Figure 3.5.2.*h*: The chaparral is dominated by shrubs. (credit: Miguel Vieira)

The Mediterranean scrub vegetation is dominated by shrubs and is adapted to periodic fires, with some plants producing seeds that germinate only after a hot fire. The ashes left behind after a fire are rich in nutrients like nitrogen and fertilize the soil, promoting plant regrowth. Fire is a natural part of the maintenance of this biome. The trees in the chaparral are mostly oaks, both deciduous and evergreen. Scrub oaks and shrubs like manzanita and the California lilac (not a relative of the eastern lilac) form dense, evergreen thickets. All of these plants are adapted to drought by such mechanisms as waxy, waterproof coatings on their leaves. Many plants that thrive in one chaparral region are successfully cultivated in other chaparral regions. For example, vineyards, olives, and figs from their native Mediterranean and eucalyptus trees from Australia thrive in California. However, blue gum eucalyptus from Australia disrupts California ecosystems by changing the hydrology and outcompeting native plants.

Temperate Grasslands

Temperate grasslands are found throughout central North America, where they are also known as prairies (figure 3.5.2.*i*), and in Eurasia, where they are known as steppes. Temperate grasslands have pronounced annual fluctuations in temperature with hot summers and cold winters. The annual temperature variation produces specific growing seasons for plants. Plant growth is possible when temperatures are warm enough to sustain plant growth, which occurs in the spring, summer, and fall.



Figure 3.5.2.*h*: The American bison (Bison bison), more commonly called the buffalo, is a grazing mammal that once populated American prairies in huge numbers. (credit: Jack Dykinga, USDA ARS)

Temperate grasslands have few trees except for those found growing along rivers or streams. The dominant vegetation tends to consist of grasses. The treeless condition is maintained by low precipitation, frequent fires, and grazing. Fires, which are a natural disturbance in temperate grasslands, can be ignited by lightning strikes. It also appears that the lightning-caused fire regime in North American grasslands was enhanced by intentional burning by humans. When fire is suppressed in temperate grasslands, the vegetation eventually converts to scrub and dense forests. Often, the restoration or management of temperate grasslands requires the use of controlled burns to suppress the growth of trees and maintain the grasses.

Annual precipitation ranges from 25.4 cm to 88.9 cm (10–35 in) and is concentrated in the summer and spring. In the plains of North America, the annual rainfall is sufficiently low (~50 cm, 20") that little or no rainfall percolates down to the water table. Calcium and other minerals are not carried below the reach of plant roots and so remain available for use. This keeps the pH and general fertility high. The grasses in undisturbed prairie are perennial. Their extensive root systems help prevent soil erosion, and the return of the season's above-ground growth to the topsoil returns minerals and provides humus to it. These advantages are diminished when annual grasses such as wheat and corn are grown instead and removed in the harvest.The self-restoring fertility of the soils of the plains states accounts for this region being the "breadbasket" of the nation (and other countries as well).

Temperate Forests

Temperate forests are the most common biome in eastern North America, Western Europe, Eastern Asia, Chile, and New Zealand (figure 3.5.2.*i*). This biome is found throughout mid-latitude regions. Temperatures range between –30°C and 30°C (–22°F to 86°F) and drop to below freezing on an annual basis. These temperatures mean that temperate forests have defined growing seasons during the spring, summer, and early fall.



Figure 3.5.2.*i*: Deciduous trees are the dominant plant in the temperate forest. (credit: Oliver Herold)

Temperate forests may be deciduous, mixed (with deciduous and evergreen trees), or coniferous (with evergreen conifers). Examples of temperate deciduous trees are beech, maple, oak, and hickory. The deciduous trees of temperate forests drop their leaves in the winter; thus, little photosynthesis occurs during the dormant winter period. (In other biomes, such as chaparral and tropical seasonal forests, trees or shrubs may drop their leaves during the summer or dry season.) Each spring, new leaves appear as temperature increases. Because of the dormant period, the net primary productivity of temperate forests is less than that of tropical rainforests. In addition, temperate forests show far less diversity of tree species than tropical rainforest biomes. Large stands dominated by a single species are common.

The deciduous trees of the temperate forests leaf out and shade much of the ground. However, more sunlight reaches the ground in this biome than in tropical rainforests because trees in temperate forests do not grow as tall as the trees in tropical rainforests. The soils of the temperate forests are rich in inorganic and organic nutrients compared to tropical rainforests. This is because of the thick layer of leaf litter on forest floors and reduced leaching of nutrients by rainfall. As this leaf litter decays, nutrients are returned to the soil. The leaf litter also protects soil from erosion, insulates the ground, and provides habitats for invertebrates and their predators. Deer, raccoons, and salamanders are characteristic inhabitants of temperate forests.

Precipitation is relatively constant throughout the year and ranges between 75 cm and 150 cm (29.5–59 in). Enough water falls on the soil so that much of it soaks deep into the ground and carries minerals with it. Such soils tend to be acidic and of low and (if unattended) diminishing fertility. Only by regular fertilization and liming (to restore calcium and raise pH) can productive agriculture be carried out in them. In the U.S., the soils east of the Appalachian Mountains tend to be of this sort.

Taiga

The **taiga**, also known as **boreal forest** in North America, is found roughly between 50° and 60° north latitude across most of Canada, Alaska, Russia, and northern Europe ((figure 3.5.2.*i*). Boreal forests are also found above a certain elevation (and below high elevations where trees cannot grow) in mountain ranges throughout the Northern Hemisphere. This biome has cold, dry winters and short, cool, wet summers. The annual precipitation is from 40 cm to 100 cm (15.7–39 in) and usually takes the form of snow. Although precipitation is relatively low, the taiga is dotted with lakes, bogs, and marshes because the cold temperatures limit evaporation.



Figure 3.5.2.*j*: The boreal forest (taiga) has low lying plants and conifer trees. (credit: L.B. Brubaker, NOAA)

The long and cold winters in the boreal forest have led to the predominance of cold-tolerant cone-bearing plants. These are evergreen coniferous trees like pines, spruce, and fir, which retain their needle-shaped leaves year-round. Evergreen trees can photosynthesize earlier in the spring than deciduous trees because less energy from the sun is required to warm a needle-like leaf than a broad leaf. In the taiga, evergreen trees grow faster than deciduous trees. In addition, soils in boreal forest regions tend to be acidic with little available nitrogen. Leaves are a nitrogen-rich structure, and deciduous trees must produce a new set of these nitrogen-rich structures each year. Therefore, coniferous trees that retain nitrogen-rich needles in a nitrogen-limiting environment may have had a competitive advantage over the broad-leafed deciduous trees.

In North America, the moose is such a typical member that it has led to the name: "spruce-moose" biome. Deer, bears, and wolves are also found in the taiga. Before the long, snowy winter sets in, many of the mammals hibernate, and many of the birds migrate south.

The net primary productivity of taiga is lower than that of temperate forests and tropical rainforests. The aboveground **biomass** (mass of living organisms) of taiga is high because these slow-growing tree species are long-lived and accumulate standing biomass over time. Species diversity is less than that seen in temperate forests and tropical rainforests. Taiga lacks the layered forest structure seen in tropical rainforests or, to a lesser degree, temperate forests. The structure of the taiga is often only a layer of trees and the understory of low-lying plants and lichen. When conifer needles are dropped, they decompose more slowly than broad leaves; therefore, fewer nutrients are returned to the soil to fuel plant growth.

At the coldest parts of the taiga, the soil remains frozen year after year, forming a layer called **permafrost**. When the permafrost melts in unusually warm years, the trees are no longer supported by this hard surface and lean in different directions, forming a "drunken forest". At extreme latitudes, the trees of the taiga become stunted by the harshness of the subarctic climate. Where they finally disappear (above the "tree line"), the Arctic tundra begins.

Arctic Tundra

The **Arctic tundra** lies north of the subarctic boreal forests and is located throughout the Arctic regions of the Northern Hemisphere. Tundra also exists at elevations above the tree line on mountains. The average winter temperature is –34°C (–29.2°F), and the average summer temperature is 3°C–12°C (37°F –52°F). Plants in the Arctic tundra have a short growing season of approximately 50–60 days. However, during this time, there are almost 24 hours of daylight and plant growth is rapid. The annual precipitation of the Arctic tundra is low (15–25 cm or 6–10 in) with little annual variation in precipitation. And, as in the taiga, there is little evaporation because of the cold temperatures.

There is little species diversity, low net primary productivity, and low above-ground biomass. Like some of the taiga, the soils of the Arctic tundra may remain in a perennially frozen state, forming permafrost. The permafrost makes it impossible for roots to penetrate far into the soil and slows the decay of organic matter, which inhibits the release of nutrients from organic matter. The surface layer of soil above the permafrost melts in the brief summer, water is released and provides for a burst of productivity while temperatures and long days permit it. Because of these harsh growing conditions, the Arctic tundra is particularly sensitive to disturbances (disruptions that decrease biomass, such as heavy machinery passing through and squishing some of the plants).

Plants in the Arctic tundra are generally low to the ground and include low shrubs, sphagnum moss, grasses, lichens, and small flowering plants (figure 3.5.2.k). Caribou feed on this growth as do vast numbers of insects. Swarms of migrating birds, especially waterfowl, invade the tundra in the summer to raise their young, feeding them on a large variety of aquatic invertebrates and vertebrates. As the brief arctic summer draws to a close, the birds fly south, and all but a few of the permanent residents, in one way or another, prepare themselves to spend the winter in a dormant state.



Figure 3.5.2.k: Low-growing plants such lichen and grasses are common in tundra. Credit: Nunavut tundra by Flickr: My Nunavut is licensed under CC-BY.

Attributions

Modified by Melissa Ha from the following sources:

- Terrestrial Biomes from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Biomes and Soil from *Biology* by John W. Kimball (licensed under CC-BY)
- Terrestrial Biomes from *Biology 2e* by OpenStax (CC-BY). Access for free at openstax.org.

This page titled 3.5.2: Terrestrial Biomes is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **8.2: Terrestrial Biomes** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

3.5.3: Aquatic Biomes

Abiotic Factors Influencing Aquatic Biomes

Like terrestrial biomes, **aquatic biomes** are influenced by a series of abiotic factors. The aquatic medium—water— has different physical and chemical properties than air. Even if the water in a pond or other body of water is perfectly clear (there are no suspended particles), water still absorbs light. As one descends into a deep body of water, there will eventually be a depth which the sunlight cannot reach. While there are some abiotic and biotic factors in a terrestrial ecosystem that might obscure light (like fog, dust, or insect swarms), usually these are not permanent features of the environment. The importance of light in aquatic biomes is central to the communities of organisms found in both freshwater and marine ecosystems. In freshwater systems, temperature stratification due to differences in density is perhaps the most critical abiotic factor and is related to the energy aspects of light. The thermal properties of water (rates of heating and cooling) are significant to the function of marine systems and have major impacts on global climate and weather patterns. Marine systems are also influenced by large-scale physical water movements, such as currents; these are less important in most freshwater lakes.

Marine Biomes

The ocean is the largest **marine biome**. It is a continuous body of salt water that is relatively uniform in chemical composition; it is a weak solution of mineral salts and decayed biological matter. Within the ocean, coral reefs are a second kind of marine biome. Estuaries, coastal areas where salt water and fresh water mix, form a third unique marine biome.

Ocean

The physical diversity of the ocean is a significant influence on plants, animals, and other organisms. The ocean is categorized by several areas (Figure 3.5.3.a). Each area has a distinct group of species adapted to the biotic and abiotic conditions particular to it. The **intertidal zone**, which is the zone between high and low tide, is the oceanic region that is closest to land. Generally, most people think of this portion of the ocean as a sandy beach. In some cases, the intertidal zone is indeed a sandy beach, but it can also be rocky or muddy. Organisms are exposed to air and sunlight at low tide and are underwater most of the time, especially during high tide. Therefore, living things that thrive in the intertidal zone are adapted to being dry for long periods of time. The shore of the intertidal zone is also repeatedly struck by waves, and the organisms found there are adapted to withstand damage from the pounding action of the waves (figure 3.5.3.b). The exoskeletons of shoreline crustaceans (such as the shore crab, *Carcinus maenas*) are tough and protect them from desiccation (drying out) and wave damage. Another consequence of the pounding waves is that few algae and plants establish themselves in the constantly moving rocks, sand, or mud.

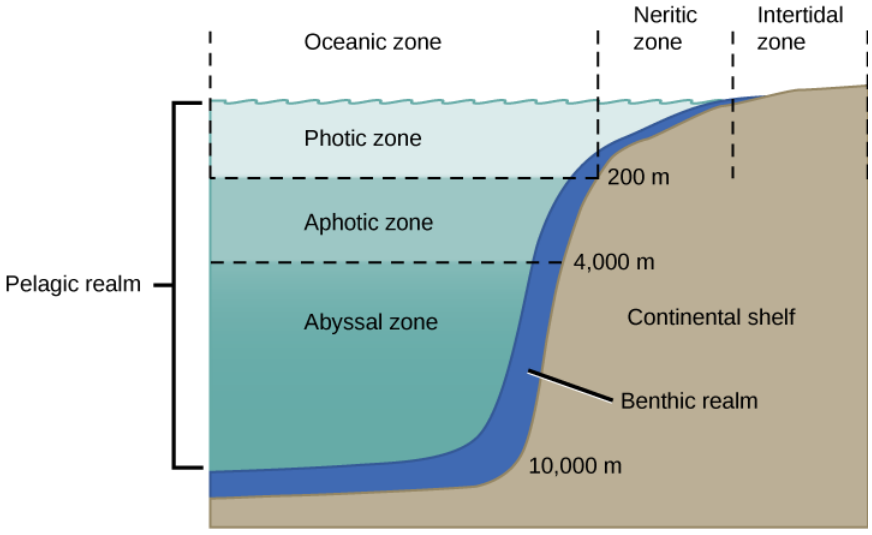


Figure 3.5.3.a: The ocean is divided into different zones based on distance from the shoreline and water depth. The intertidal zone is the closest to shore followed by the neretic and oceanic zone. The photic zone is 0-200 meters deep. The aphotic zone is 200-4,000 meters deep. The abyssal zone is 4,000-10,000 meters deep. These three zones form the pelagic realm, and the benthic realm is below the pelagic realm, along the continental shelf.



Figure 3.5.3.*b*: Sea urchins, mussel shells, and starfish are often found in the intertidal zone, shown here in Kachemak Bay, Alaska. (credit: NOAA)

The **neritic zone** extends from the intertidal zone to depths of about 200 m (or 650 ft) at the edge of the continental shelf. Because light can penetrate this depth, photosynthesis can occur. The water here contains silt and is well-oxygenated, low in pressure, and stable in temperature. Phytoplankton and floating *Sargassum* (a type of free-floating marine seaweed) provide a habitat for some sea life found in the neritic zone. Zooplankton, protists, small fishes, and shrimp are found in the neritic zone and are the base of the food chain for most of the world's fisheries.

Beyond the neritic zone is the open ocean area known as the **oceanic zone**. Within the oceanic zone there is thermal stratification where warm and cold waters mix because of ocean currents. Abundant plankton serve as the base of the food chain for larger animals such as whales and dolphins. Nutrients are scarce and this is a relatively less productive part of the marine biome. When photosynthetic organisms and the protists and animals that feed on them die, their bodies fall to the bottom of the ocean where they remain.

All of the ocean's open water is referred to as the **pelagic realm**. The pelagic realm is divided into the photic, aphotic, and abyssal zones from top to bottom based on how far light reaches into the water. The **photic zone**, which is the portion of the ocean that light can penetrate (approximately 200 m or 650 ft). At depths greater than 200 m, light cannot penetrate; thus, this is referred to as the **aphotic zone**. The majority of organisms in the aphotic zone include sea cucumbers (phylum Echinodermata) and other organisms that survive on the nutrients contained in the dead bodies of organisms in the photic zone.

The deepest part of the ocean is the **abyssal zone**, which is at depths of 4000 m or greater. Both the aphotic and abyssal zones lack sufficient light for photosynthesis, and together, they constitute most of the ocean. The deepest part of the ocean, the Challenger Deep (in the Mariana Trench, located in the western Pacific Ocean), is about 11,000 m (about 6.8 mi) deep. To give some perspective on the depth of this trench, the ocean is, on average, 4267 m deep. These zones are relevant to freshwater lakes as well. The abyssal zone is very cold and has very high pressure, high oxygen content, and low nutrient content. There are a variety of invertebrates and fishes found in this zone, but the abyssal zone does not have plants because of the lack of light. Cracks in the Earth's crust called **hydrothermal vents** are found primarily in the abyssal zone (figure 3.5.3.*c*). Around these vents, bacteria that utilize the hydrogen sulfide and other minerals emitted as an energy source serve as the base of the food chain found in the abyssal zone.

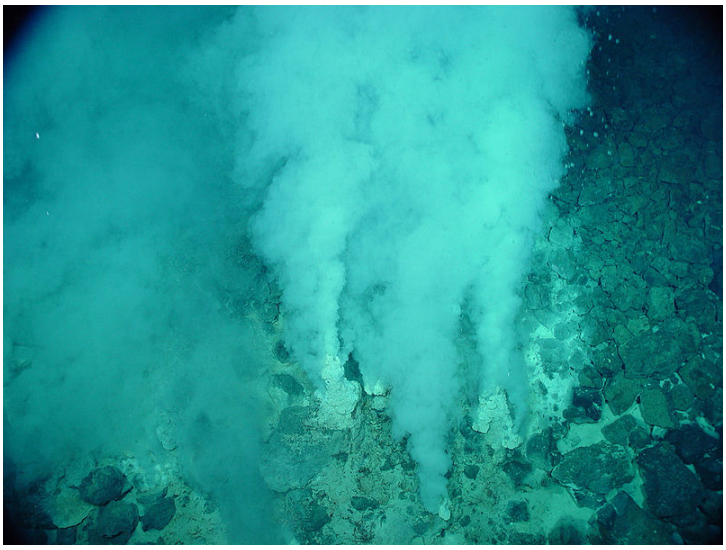


Figure 3.5.3.*c*: A hydrothermal vent. Image by NOAA (public domain).

The **benthic realm**, extends along the ocean bottom from the shoreline to the deepest parts of the ocean floor. It is comprised of sand, silt, and dead organisms. This is a nutrient-rich portion of the ocean because of the dead organisms that fall from the upper layers of the ocean. Because of this high level of nutrients, a diversity of sponges, sea anemones, marine worms, sea stars, fishes, and bacteria exist.

Coral Reefs

Coral reefs are characterized by high biodiversity and the structures created by invertebrates that live in warm, shallow waters within the photic zone of the ocean. They are mostly found within 30 degrees north and south of the equator. The Great Barrier Reef is a well-known reef system located several miles off the northeastern coast of Australia. The coral organisms are colonies of saltwater polyps that secrete a calcium carbonate skeleton. These calcium-rich skeletons slowly accumulate, forming the underwater reef (figure 3.5.3.*d*).



Figure 3.5.3.*d*: Coral reefs are formed by the calcium carbonate skeletons of coral organisms, which are marine invertebrates. (credit: Terry Hughes)

Corals found in shallower waters (at a depth of approximately 60 m or about 200 ft) have a mutualistic relationship with photosynthetic unicellular algae called dinoflagellates. The relationship provides corals with the majority of the nutrition and the energy they require. The waters in which these corals live are nutritionally poor, and, without this mutualism, it would not be possible for large corals to grow. Some corals living in deeper and colder water do not have a mutualistic relationship with algae; these corals attain energy and nutrients using stinging cells on their tentacles to capture prey. It is estimated that more than 4,000 fish species inhabit coral reefs. These fishes can feed on coral, other invertebrates, or the seaweed and algae that are associated with the coral.

Estuaries: Where the Ocean Meets Fresh Water

Estuaries are biomes that occur where a source of fresh water, such as a river, meets the ocean. Therefore, both fresh water and salt water are found in the same vicinity; mixing results in a diluted (brackish) saltwater. Estuaries form protected areas where many of the young offspring of crustaceans, mollusks, and fish begin their lives. Salinity is a very important factor that influences the organisms and the adaptations of the organisms found in estuaries. The salinity of estuaries varies and is based on the rate of flow of its freshwater sources. Once or twice a day, high tides bring salt water into the estuary. Low tides occurring at the same frequency reverse the current of salt water.

The short-term and rapid variation in salinity due to the mixing of fresh water and salt water is a difficult physiological challenge for the plants and animals that inhabit estuaries. Many estuarine plant species are **halophytes**, plants that can tolerate salty conditions. Halophytic plants are adapted to deal with the salinity resulting from saltwater on their roots or from sea spray. In some halophytes, filters in the roots remove the salt from the water that the plant absorbs. Other plants are able to pump oxygen into their roots. Animals, such as mussels and clams, have developed behavioral adaptations that expend a lot of energy to function in this rapidly changing environment. When these animals are exposed to low salinity, they stop feeding, close their shells, and stop using oxygen. When high tide returns to the estuary, the salinity and oxygen content of the water increases, and these animals open their shells, begin feeding, and return to using oxygen.

Freshwater Biomes

Freshwater biomes include lakes and ponds (standing water) as well as rivers and streams (flowing water). They also include wetlands, which will be discussed later. Humans rely on freshwater biomes to provide aquatic resources for drinking water, crop irrigation, sanitation, and industry. These various roles and human benefits are referred to as ecosystem services. Lakes and ponds are found in terrestrial landscapes and are, therefore, connected with abiotic and biotic factors influencing these terrestrial biomes.

Lakes and Ponds

Lakes and ponds can range in area from a few square meters to thousands of square kilometers. Temperature is an important abiotic factor affecting living things found in lakes and ponds. In the summer, thermal stratification of lakes and ponds occurs when the upper layer of water is warmed by the sun and does not mix with deeper, cooler water. Like oceans, lakes and ponds have photoic zones through which light can penetrate and aphotic zones without light. **Phytoplankton** (small photosynthetic organisms such as algae and photosynthetic bacteria that float in the water) are found here and carry out photosynthesis, providing the base of the food web of lakes and ponds. **Zooplankton** (very small animals that float in the water), such as rotifers and small crustaceans, consume these phytoplankton (figure 3.5.3.e). At the bottom of lakes and ponds, bacteria in the aphotic zone break down dead organisms that sink to the bottom.

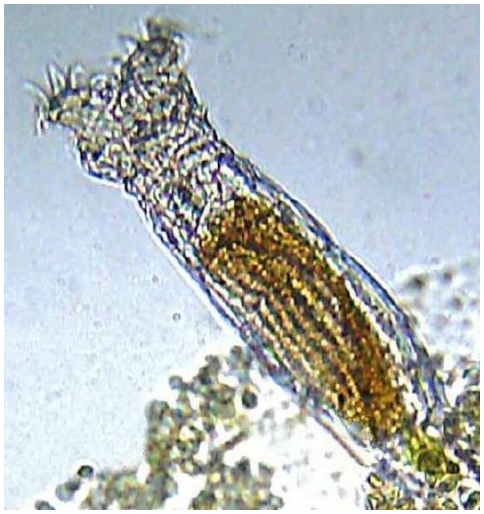


Figure 3.5.3.e: A rotifer is an example of the zooplankton found in lakes and ponds. Image by Bob Blaylock at English Wikipedia (CC-BY-SA).

Rivers and Streams

Rivers and streams are continuously moving bodies of water that carry large amounts of water from the source, or headwater, to a lake or ocean. The largest rivers include the Nile River in Africa, the Amazon River in South America (Figure 3.5.3.f), and the Mississippi River in North America. Abiotic features of rivers and streams vary along the length of the river or stream. Streams begin at a point of origin referred to as source water. The source water is usually cold, low in nutrients, and clear. The channel (the width of the river or stream) is narrower than at any other place along the length of the river or stream. Because of this, the current is often faster here than at any other point of the river or stream.



Figure 3.5.3.*f*: The Amazon River is the world's largest river. Image by Jason Hollinger (CC-BY)

The fast-moving water results in minimal silt accumulation at the bottom of the river or stream, therefore the water is clear. Photosynthesis here is mostly attributed to algae that are growing on rocks; the swift current inhibits the growth of phytoplankton. An additional input of energy can come from leaves or other organic material that falls into the river or stream from trees and other plants that border the water. When the leaves decompose, the organic material and nutrients in the leaves are returned to the water. Plants and animals have adapted to this fast-moving water. For instance, leeches have elongated bodies and suckers on both ends. These suckers attach to the substrate, keeping the leech anchored in place. Freshwater trout species are an important predator in these fast-moving rivers and streams.

As the river or stream flows away from the source, the width of the channel gradually widens and the current slows. This slow-moving water, caused by the gradient decrease and the volume increase as tributaries unite, has more sedimentation. Phytoplankton can also be suspended in slow-moving water. Therefore, the water will not be as clear as it is near the source. The water is also warmer. Worms and insects can be found burrowing into the mud. The higher order predator vertebrates include waterfowl, frogs, and fishes.

Wetlands

Wetlands are environments in which the soil is either permanently or periodically saturated with water. Wetlands are different from lakes because wetlands are shallow bodies of water that may periodically dry out. Emergent vegetation consists of wetland plants that are rooted in the soil but have portions of leaves, stems, and flowers extending above the water's surface. There are several types of wetlands including marshes, swamps, bogs, mudflats, and salt marshes (figure 3.5.3.*g*).



Figure 3.5.3.*g*: Located in southern Florida, Everglades National Park is vast array of wetland environments, including sawgrass marshes, cypress swamps, and estuarine mangrove forests. Here, a Great Egret walks among cypress trees. (credit: NPS)

Attribution

Modified by Melissa Ha from Aquatic Biomes from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 3.5.3: Aquatic Biomes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **8.3: Aquatic Biomes** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

3.5.4: Data Dive- Biome Carbon Storage

Overview

Carbon sequestration is the process by which atmospheric carbon dioxide is taken up by trees, grasses, and other plants through photosynthesis and stored as carbon in biomass (trunks, branches, foliage, and roots). Carbon storage can also occur in soil formed from organic matter found in decomposing litter. Terrestrial carbon sequestration has been a recent topic of interest because of its implications for use in climate change mitigation. The carbon storage in terrestrial biomes helps to offset sources of carbon dioxide to the atmosphere, such as deforestation, wildfires, and fossil fuel emissions. As such, carbon storage in each biome can be altered by both natural and human caused events. The figure below indicates the estimated storage above (plants) and below (soil) ground by terrestrial biomes:

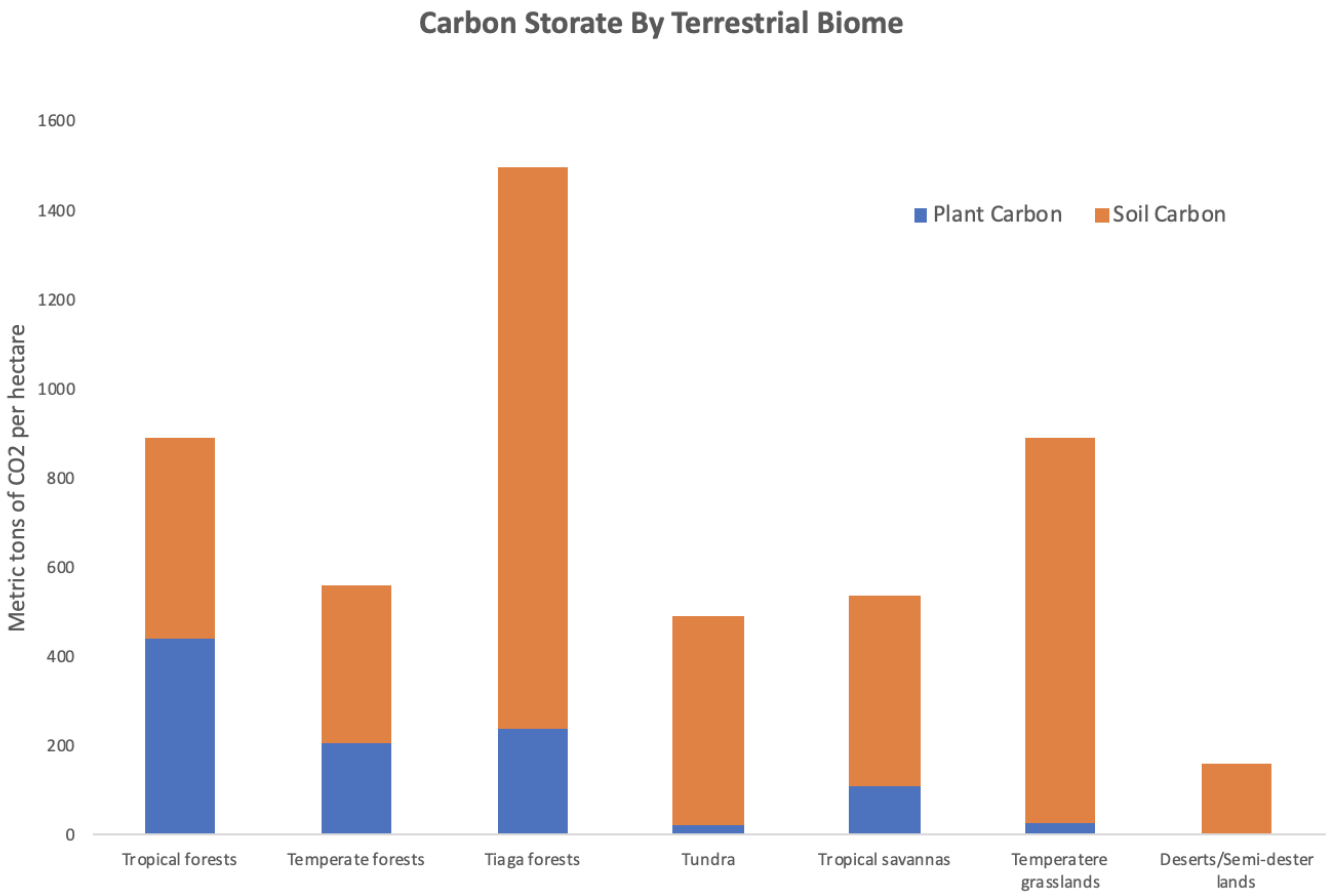


Figure 3.5.4.*a*: Carbon storage (in metric tons of CO2 per hectare) above (plant) and below (soil) ground in various biomes. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Islam KK, Sato N 2012.

Questions

1. What kind of graph is this?
2. What is the independent (explanatory) variable and the dependent (response) variable?
3. Which biome has the greatest storage in plants? What about soil?
4. What are the top three terrestrial biomes in terms of their carbon storage?
5. How can the results of this graph to inform future conservation efforts?
6. What do the results of this graph make you curious about?

Raw Data From Above Graph(s)

Table 3.5.4.*a*: Raw data of carbon storage (in metric tons of CO2 per hectare) above (plant) and below (soil) ground in various biomes. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Islam KK, Sato N 2012.

Biome	Plant Carbon	Soil Carbon	Total Carbon
Tropical Forests	442	450	892
Temperate Forests	208	352	561
Taiga Forests	236	1260	1496
Tundra	23	467	490
Tropical Savannas	108	430	538
Temperate Grasslands	26	865	892
Deserts/Semi-Desert Lands	6	154	160

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 3.5.4: Data Dive- Biome Carbon Storage is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **8.4: Data Dive- Biome Carbon Storage** by Melissa Ha and Rachel Schleiger has no license indicated.

3.5.5: Review

Summary

After completing this chapter you should be able to...

- Describe how latitude and altitude affect temperature and precipitation, which in turn, determine the distribution of biomes.
- Compare the eight major terrestrial biomes.
- Describe how light, temperature, flow, and salinity influence aquatic biomes.
- Distinguish between marine and freshwater biomes and describe examples of each.
- Explain the different realms and zones of the ocean.

Earth has **terrestrial biomes** and **aquatic biomes**. Temperature and precipitation, and variations in both, are key abiotic factors that shape the composition of animal and plant communities in terrestrial biomes. There are eight major terrestrial biomes: **tropical rainforests, savannas, deserts, chaparral, temperate grasslands, temperate forests, taiga,** and **Arctic tundra**.

Like terrestrial biomes, aquatic biomes are influenced by abiotic factors. In the case of aquatic biomes, the abiotic factors include light, temperature, flow regime, and dissolved solids. Aquatic biomes include both **marine biomes** and **freshwater biomes**. **Oceans**, **coral reefs**, and **estuaries** are marine biomes while **lakes**, **ponds**, **rivers**, **streams**, and **wetlands** are freshwater biomes.

Attribution

Modifed by Melissa Ha from Ecosystems and the Biosphere- Chapter Resources from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 3.5.5: Review is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **8.5: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY 4.0.

CHAPTER OVERVIEW

4: Conservation

Conservation biology is the management of Earth's ecosystems with the aim of protecting species, their communities, and ecosystems from escalated extinction rates and the destruction and degradation of their habitats. This field focuses on both evaluating past, current, and future trends and creating conservation action plans to reverse trends leading to extinction and destroyed habitats. More specifically, procedures have been put in place to protect threatened species, design habitat preserves, creating breeding programs, and reconciling conservation concerns with practical human needs.

As conservation biology needs to pull from so many fields, it is an interdisciplinary subject drawing on physical, life, and social sciences, as well as natural resource management procedures. The field seeks to integrate social science policy with theories from the fields of ecology, demography, taxonomy, and genetics. The principles underlying each of these disciplines have direct implications for the management of species and ecosystems, captive breeding and reintroduction, genetic analyses, and habitat restoration. Social science disciplines not only help with putting policies in place to legally protect species and habitats but also help fund conservation actions. Conservation action plans help direct research, monitoring, and education programs that engage concerns at both local and global scales.

The goal of this unit is to overview the essential conservation biology topics as we head into an uncertain future.



Figure 4.*a*: A road heading into our uncertain future. Image by Pikist (Public domain)

Attributions

Rachel Schleiger (CC-BY-NC)

4.1: The Value of Biodiversity

4.1.1: Ecosystem Diversity

4.1.2: Species Diversity

4.1.3: Genetic Diversity

4.1.4: Patterns of Biodiversity

4.1.5: Data Dive- Biodiversity and Drugs

4.1.6: Review

4.2: Threats to Biodiversity

4.2.1: Extinction

4.2.2: Measures of Biodiversity Loss

4.2.3: Habitat Loss

4.2.4: Overexploitation

4.2.5: Pollution
4.2.6: Invasive Species
4.2.7: Climate Change
4.2.8: Data Dive- Threats to Biodiversity
4.2.9: Review
4.3: Protecting Biodiveristy
4.3.1: Policies
4.3.2: Non-profit Organizations
4.3.3: Species-level Conservation
4.3.4: Protected Areas
4.3.5: Ecosystem Restoration
4.3.6: Economic Influences on Conservation
4.3.7: Individual Choices
4.3.8: Data Dive- Island Fox Populations
4.3.9: Review

This page titled 4: Conservation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

SECTION OVERVIEW

4.1: The Value of Biodiversity

Chapter Hook

Most of the medicine we depend on today originated because of the diversity of plants across the globe. New medicines sometimes come from newly found species, but there is still much to learn from already discovered species as well. In 2019, research was published about a newly discovered quality on a well-known California plant called yerba santa (*Eriodictyon californicum*). Among native people of California, yerba santa was already well known to have medicinal qualities, but only recently did research scientists begin to test them. It was discovered that a compound in yerba santa decreases the onset of Alzheimer's disease, a leading cause of death in the United States, by having protective and supportive qualities on the neurons in our nervous system. This research is not only vital from a medical perspective, but also exemplifies the importance of protecting diversity.



Figure 4.1.*a* Yerba santa (*Eriodictyon californicum*) in bloom. Image by Breck22 (Public domain)

Biodiversity is a broad term for the variety of life on Earth. Traditionally, ecologists have measured biodiversity by taking into account both the number of species and the number of individuals of each species. However, biologists now measure biodiversity at a number of organizational levels, including ecosystem, species, and genetic diversity. This focuses efforts to preserve the biologically and technologically important elements of biodiversity. Biodiversity is important to the survival and welfare of human

populations because it has impacts on our health and our ability to feed ourselves through agriculture and harvesting populations of wild animals.

Attribution

Modified by Rachel Schleiger and Melissa Ha from Importance of Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

4.1.1: Ecosystem Diversity

4.1.2: Species Diversity

4.1.3: Genetic Diversity

4.1.4: Patterns of Biodiversity

4.1.5: Data Dive- Biodiversity and Drugs

4.1.6: Review

This page titled 4.1: The Value of Biodiversity is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

4.1.1: Ecosystem Diversity

Measuring biodiversity on a large scale involves measuring **ecosystem diversity**, the number of different ecosystems on Earth or in a geographical area as well as their relative abundances (figure 4.1.1.a). The loss of an ecosystem means the loss of the interactions between species and the loss of biological productivity that an ecosystem is able to create. An example of a largely extinct ecosystem in North America is the prairie ecosystem (figure 4.1.1.a). Prairies once spanned central North America from the boreal forest in northern Canada down into Mexico. They are now all but gone, replaced by crop fields, pasture lands, and suburban sprawl. Many of the species survive, but the hugely productive ecosystem that was responsible for creating our most productive agricultural soils is now gone. As a consequence, their soils are now being depleted unless they are maintained artificially at great expense.



Figure 4.1.1.a: The variety of ecosystems on Earth—from the coral reef (top) to prairie (bottom)—enables a great diversity of species to exist. Top image by Wise Hok Wai Lum (CC-BY-SA). Bottom image modified from Jim Minnerath, USFWS.

The soil productivity described above is an example of an **ecosystem service**. These are the products and processes associated with biological systems and are directly or indirectly of immense value to the well-being of people. Some ecosystem services are processes such as the regulation of climate, flooding, and disease. Nutrient cycling, pollination, and regulation of crop pests are ecosystem services important to food production (see the Food Production and Ecosystem Services box below). The water cycle provides fresh water, and photosynthesis adds oxygen to our air. Other ecosystem services are human provisions including food, fuel, and fiber (such as cotton for clothing or timber). Medicines are another important provision (see Importance of Species Diversity). Furthermore, healthy ecosystems allow for recreational activities, such as hiking, kayaking, and camping, and educational opportunities, such as field trips. Nature is also the basis for a significant part of aesthetic and spiritual values held by many cultures.

In 1997, Robert Costanza and his colleagues estimated to annual value of ecosystem services to be $33 trillion dollars ($53 trillion in 2019 dollars), and many consider this to be an underestimation. For comparison the gross domestic product of the United States

in 2020 was $21 trillion. Valuing ecosystem services can be difficult, particular for those services that are processes rather than provisions. One strategy is to calculate replacement cost. For example, how much would it cost to control a pest population if it was not regulated by natural processes? Figure 4.1.1.*b* illustrates the value of a few ecosystem services.

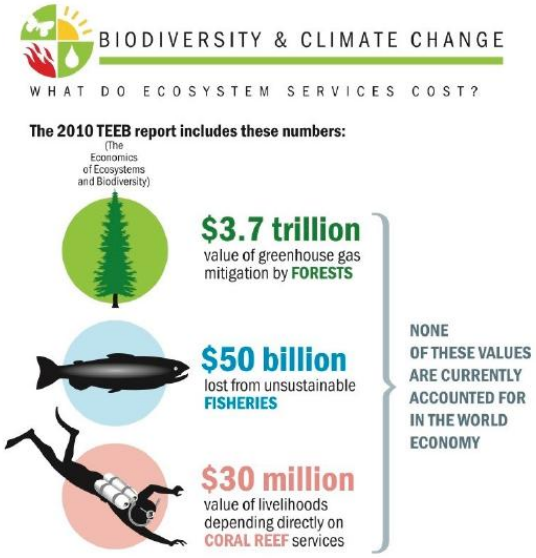


Figure 4.1.1.*b*: The value of three ecosystem services. The question, "What do ecosystem services cost?" is partially answered by the 2010 TEEB (The Economics of Ecosystems and Biodiversity) report. Greenhouse gas mitigation (which limits climate change) provided by forests is valued at $3.7 trillion. Unsustainable fisheries have compromised fish availability, costing $50 billion. Coral reef services, such as tourism, are valued at $30 million. Values are in 2010 dollars. None of these values are currently accounted for in the world economy. Data from The Economics of Ecosystems and Biodiversity Synthesis Report. Image by Future Challenges (CC-BY-SA).

Food Production and Ecosystem Services

Food production relies on several interacting ecosystem services. Most soils contain a huge diversity of organisms that maintain nutrient cycles, breaking down organic matter into nutrient compounds that crops need for growth. These organisms also maintain soil texture that affects water and oxygen dynamics in the soil that are necessary for plant growth. Replacing the work of these organisms in forming **arable** (farmable) soil is not practically possible. They occur within ecosystems, such as soil ecosystems, as a result of the diverse metabolic activities of the organisms living there, but they provide benefits to human food production, drinking water availability, and breathable air.

Other key ecosystem services related to food production are plant pollination and crop pest control. It is estimated that honeybee pollination within the United States brings in $1.6 billion per year; other pollinators contribute up to $6.7 billion. Over 150 crops in the United States require pollination to produce. Many honeybee populations are managed by beekeepers who rent out their hives' services to farmers. Honeybee populations in North America have been suffering large losses caused by a syndrome known as **colony collapse disorder**, a phenomenon with complex and interacting causes.

Other pollinators include a diverse array of other bee species and various insects and birds. Loss of these species would make growing crops requiring pollination impossible, increasing dependence on other crops. A 2002 study by Claire Kremen and colleagues found that **native** pollinators in Central California (those that historically occurred there; figure 4.1.1.*c* and d) provided full pollination to watermelon crops. This was only true on organic farms that were located near the natural habitat for these pollinators, highlighting the importance of sustainable farming practices and habitat conservation in preserving ecosystem services. Essentially, in a healthy ecosystem, native pollinators eliminated the need for rented honeybee hives. Note that while honeybees are valuable in an agricultural setting, they are not native to North America and can disrupt ecosystems by competing with native bees.



Figure 4.1.1.*c*: The California bumble bee (*Bombus californicus*) is one of several native insects that provided watermelon pollination in Kremen and colleagues' study. Image by Rhododendrites (CC-BY-SA).



Figure 4.1.1.*d*: Mason bees are efficient pollinators, and many are native to North America. Image by Robert Engelhardt (CC-BY-SA).

Finally, humans compete for their food with crop pests, most of which are insects. Pesticides control these competitors, but these are costly and lose their effectiveness over time as pest populations adapt. They also lead to collateral damage by killing non-pest species as well as beneficial insects like honeybees, and risking the health of agricultural workers and consumers. Moreover, these pesticides may migrate from the fields where they are applied and do damage to other ecosystems like streams, lakes, and even the ocean (see Industrial Agriculture and Environmental Toxicology). Ecologists believe that the bulk of the work in removing pests is actually done by predators and parasites of those pests, but the impact has not been well studied. A review found that in 74 percent of studies that looked for an effect of landscape complexity (forests and fallow fields near to crop fields) on natural enemies of pests, the greater the complexity, the greater the effect of pest-suppressing organisms. Another experimental study found that introducing multiple enemies of pea aphids (an important alfalfa pest) increased the yield of alfalfa significantly. This study shows that a diversity of pests is more effective at control than one single pest. Loss of diversity in pest enemies will inevitably make it more difficult and costly to grow food. The world's growing human population faces significant challenges in the increasing costs and other difficulties associated with producing food.

References

Costanza, R., d'Arge, R., de Groot, R. *et al.* The value of the world's ecosystem services and natural capital. *Nature* **387,** 253–260 (1997). DOI

Kremen, C., Williams, N. M., and Thorp, R. W. Crop pollination from native bees at risk from agricultural intensification. *PNAS* **99,** 6812-16816 (2002). DOI

Attributions

Modified by Melissa Ha from the following sources:

- Importance of Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Biodiversity, Species Loss, and Ecosystem Function from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors (licensed under CC-BY). Download for free at CNX.

This page titled 4.1.1: Ecosystem Diversity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **9.1: Ecosystem Diversity** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

4.1.2: Species Diversity

Species diversity is the number of different species in a particular area and their relative abundance. The area in question could be a habitat, a biome, or the entire biosphere. Areas with low species diversity, such as the glaciers of Antarctica, still contain a wide variety of living organisms, whereas the diversity of tropical rainforests is so great that it cannot be accurately assessed. **Species richness**, the number of species living in a habitat or other unit, is one component of biodiversity. **Species evenness** is a component of species diversity based on **relative abundance** (the number individuals in a species relative to the total number of individuals in all species within a system). Foundation species (see Ecosystem Types and Dynamics) often have the highest relative abundance of species. Two locations with the same richness do not necessarily have the same species evenness. For example, both communities in figure 4.1.2.*b* have three different trees species and thus a species richness of three. However, there is a dominant species (represented by six individuals) in community #1. In community #2, there are three of individuals of each species. Therefore, community #2 has a greater species evenness and greater species diversity overall.

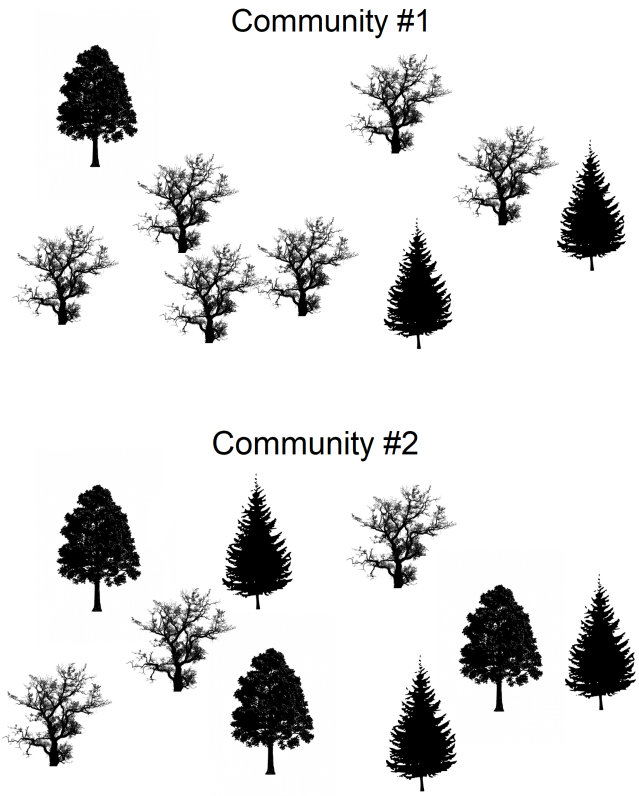


Figure 4.1.2.*a*: Two hypothetical tree communities have the same species richness, but community #2 (bottom) has a greater species evenness. Both communities have nine trees and three tree species. In community #1, one species is dominant, represented by six individuals. There are two individuals of a conifer species, and only one individual of the final species. In community #2, there are three individuals from each species. Images compiled by Melissa Ha from Alone tree George Hodan, Old Tree Silhouette, and Tree (all public domain).

The Number of Species on Earth

Despite considerable effort, knowledge of the species that inhabit the planet is limited. About 1.5 million species have been described, but many more species are yet to be identified. Estimates of the total number of species on Earth range from 3 million to 100 million, with more recent estimates commonly ranging from 8 to 11 million species. A 2011 study suggests that only 13% of **eukaryotic** species (such as plants, animals, fungi, and algae) have been named (Table 4.1.2.*a* . Estimates of numbers of **prokaryotic** species (such as bacteria) are largely guesses, but biologists agree that science has only just begun to catalog their diversity. In fact, a 2017 study by Brendan Larsen and colleagues estimated that there are actually 1-6 billion species on Earth with

at least 70% of them being bacteria. Given that Earth is losing species at an accelerating pace, science knows little about what is being lost.

Table 4.1.2.*a***:** The estimated number of species by taxonomic group—including both described (named and studied) and predicted (yet to be named) species.

Type of Organism	Mora et al. 2011 Described	Mora et al. 2011 Predicted	Chapman 2009 Described	Chapman 2009 Predicted	Groombridge and Jenkins 2002 Described	Groombridge and Jenkins 2002 Predicted
Animals	1,124,516	9,920,000	1,424,153	6,836,330	1,225,500	10,820,000
Photosynthetic protists (such as algae)	17,892	34,900	25,044	200,500		
Fungi	44,368	616,320	98,998	1,500,000	72,000	1,500,000
Plants	224,244	314,600	310,129	390,800	270,000	320,000
Non-photosynthetic protists	16,236	72,800	28,871	1,000,000	80,000	600,000
Prokaryotes			10,307	1,000,000	10,175	
Total	1,438,769	10,960,000	1,897,502	10,897,630	1,657,675	13,240,000

There are various initiatives to catalog described species in accessible and more organized ways, and the internet is facilitating that effort. Nevertheless, at the current rate of species description, which according to the State of Observed Species reports is 17,000–20,000 new species a year, it would take close to 500 years to describe all of the species currently in existence. The task, however, is becoming increasingly impossible over time as extinction removes species from Earth faster than they can be described.

Naming and counting species may seem an unimportant pursuit given the other needs of humanity, but it is not simply an accounting. Describing species is a complex process by which biologists determine an organism's unique characteristics and whether or not that organism belongs to any other described species. It allows biologists to find and recognize the species after the initial discovery to follow up on questions about its biology. That subsequent research will produce the discoveries that make the species valuable to humans and to our ecosystems. Without a name and description, a species cannot be studied in depth and in a coordinated way by multiple scientists.

DNA Barcoding

The technology of molecular genetics and data processing and storage are maturing to the point where cataloging the planet's species in an accessible way is close to feasible. **DNA barcoding** is one molecular genetic method, which takes advantage of a specialized structures inside of some cells called mitochondria (figure 4.1.2.*a*). **Mitochondria** contain DNA that is separate from the rest of the cell, and one of the genes in mitochondrial DNA changes more quickly through the process of evolution than regular DNA. While plants contain mitochondria, DNA from their **chloroplasts**, the specialized structures in which photosynthesis occurs, are more often barcoded. Rapid DNA sequencing technology make the molecular genetics portion of the work relatively inexpensive and quick. Computer resources store and make available the large volumes of data. Projects are currently underway to use DNA barcoding to catalog museum specimens, which have already been named and studied, as well as testing the method on less studied groups. As of mid 2012, close to 150,000 named species had been barcoded. Early studies suggest there are significant numbers of undescribed species that looked too much like sibling species to previously be recognized as different. These now can be identified with DNA barcoding.

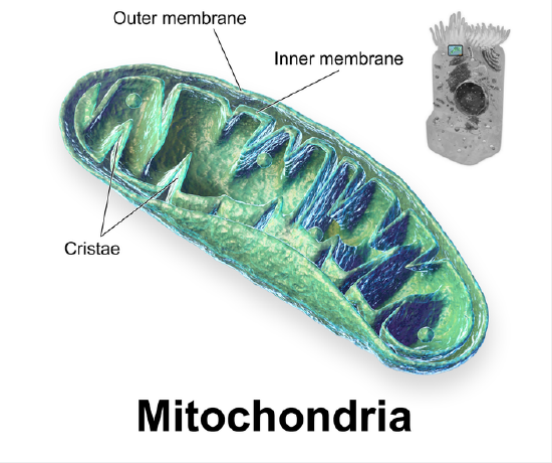


Figure 4.1.2.*b*: Mitochondria are specialized structures inside of eukaryotic cells. They contain separate DNA, and mitochondrial DNA contains a rapidly evolving gene used for DNA barcoding. Image by Blausen.com staff (2014). "Medical gallery of Blausen Medical 2014". *WikiJournal of Medicine* **1** (2). DOI:10.15347/wjm/2014.010. ISSN 2002-4436. (CC-BY)

Importance of Species Diversity

Healthy ecosystems contain a diversity of species, and each species plays a role in ecosystem function; therefore, species diversity as well as ecosystem diversity are essential to maintaining ecosystem services. For example, many medications are derived from natural chemicals made by a diverse group of organisms. For example, many plants produce compounds meant to protect the plant from insects and other animals that eat them. Some of these compounds also work as human medicines. Contemporary societies that live close to the land often have a broad knowledge of the medicinal uses of plants growing in their area. For centuries in Europe, older knowledge about the medical uses of plants was compiled in herbals—books that identified the plants and their uses. Humans are not the only animals to use plants for medicinal reasons. The other great apes, orangutans, chimpanzees, bonobos, and gorillas have all been observed self-medicating with plants.

Modern pharmaceutical science also recognizes the importance of these plant compounds. Examples of significant medicines derived from plant compounds include aspirin, codeine, digoxin, atropine, and vincristine (figure 4.1.2.*c*). Many medications were once derived from plant extracts but are now synthesized. It is estimated that, at one time, 25 percent of modern drugs contained at least one plant extract. That number has probably decreased to about 10 percent as natural plant ingredients are replaced by synthetic versions of the plant compounds. Antibiotics, which are responsible for extraordinary improvements in health and lifespans in developed countries, are compounds largely derived from fungi and bacteria.



Figure 4.1.2.*c*: *Catharanthus roseus*, the Madagascar periwinkle, has various medicinal properties. Among other uses, it is a source of vincristine, a drug used in the treatment of lymphomas. (credit: Forest and Kim Starr)

In recent years, animal venoms and poisons have excited intense research for their medicinal potential. By 2007, the FDA had approved five drugs based on animal toxins to treat diseases such as hypertension, chronic pain, and diabetes. Another five drugs

are undergoing clinical trials and at least six drugs are being used in other countries. Other toxins under investigation come from mammals, snakes, lizards, various amphibians, fish, snails, octopuses, and scorpions.

Aside from representing billions of dollars in profits, these medications improve people's lives. Pharmaceutical companies are actively looking for new natural compounds that can function as medicines. It is estimated that one third of pharmaceutical research and development is spent on natural compounds and that about 35 percent of new drugs brought to market between 1981 and 2002 were from natural compounds.

References

Brendan B. Larsen, Elizabeth C. Miller, Matthew K. Rhodes, and John J. Wiens, "Inordinate Fondness Multiplied and Redistributed: the Number of Species on Earth and the New Pie of Life," The Quarterly Review of Biology 92, no. 3 (September 2017): 229-265. DOI

Attributions

Modified by Melissa Ha from the following sources:

- Community Ecology and Importance of Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Preserving Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)

This page titled 4.1.2: Species Diversity is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **9.2: Species Diversity** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

4.1.3: Genetic Diversity

In contrast to ecosystem and species diversity, **genetic diversity** is a measure of the variability among individuals within a single species. Genetic diversity is represented by the variety of alleles present within a population. A **gene** represents the fundamental physical unit of heredity, and **alleles** are specific versions of these genes. For example, there is a gene for flower colors in peas, one allele for that gene carries information for white flowers while another allele carries information for purple flowers. Alleles are passed from parent to offspring. **Mutations**, changes in the DNA sequences that usually result from errors during DNA replication, are the original source of new alleles.

As a result of sexual reproduction, individuals can have different allele combinations for each gene. Genetic diversity provides the raw material for **evolutionary adaptation**, the process by which the genetic composition of populations change over time in a way that makes it more suited to the environment.

Loss of genetic diversity makes a species less able to reproduce successfully and less adaptable to a changing environment or to a new disease. Small populations of species are especially susceptible to loss of genetic diversity. When a species loses too many individuals, it becomes genetically uniform. Some of the causes for the loss in genetic diversity include: **inbreeding** among closely related individuals and **genetic drift**, the process by which the genetic composition of a population fluctuates randomly over time.

Low genetic diversity makes the Tasmanian devil (figure 4.1.3.a) especially vulnerable to Devil Facial Tumor Disease (DFTD), a cancer that threatens it with extinction. Contagious cancers are typically spread by viruses, which can transmit cancer-causing genes; however, in the case of DFTD, the cancerous cells themselves are spread among individuals. Usually, the immune system can recognize cancerous cells, infected cells, and foreign cells, but DFTD cells evade the immune system. If Tasmanian devils had high genetic diversity, it is likely that some individuals would have alleles that made them resistant to DFTD. These individuals would survive and reproduce more frequently than others, and the resistant alleles would become more common in the population (evolutionary adaptation would occur). Because Tasmanian devils have low genetic diversity, there is less opportunity for such evolutionary adaptation.

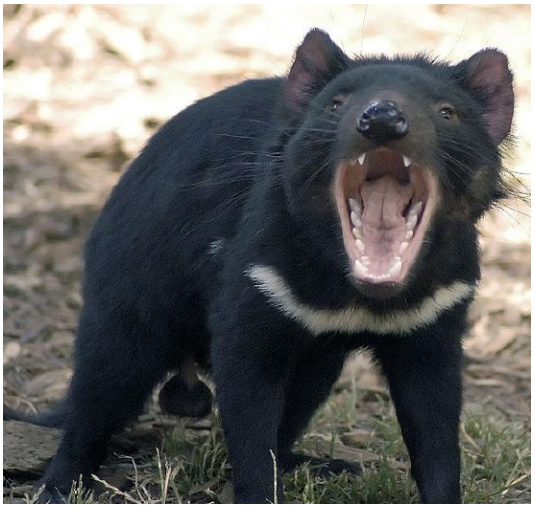


Figure 4.1.3.a: The Tasmanian devil in a defensive stance, at Tasmanian Devil Conservation Park, Tasman Peninsula. Image by Wayne McLean (CC-BY-SA).

Genetic diversity is important to agriculture. Since the beginning of human agriculture more than 10,000 years ago, human groups have been breeding and selecting crop varieties. This crop diversity matched the cultural diversity of highly subdivided populations of humans. For example, potatoes were domesticated beginning around 7,000 years ago in the central Andes of Peru and Bolivia. The people in this region traditionally lived in relatively isolated settlements separated by mountains. The potatoes grown in that region belong to seven species and the number of varieties likely is in the thousands. Each variety has been bred to thrive at particular elevations and soil and climate conditions. The diversity is driven by the diverse demands of the dramatic elevation changes, the limited movement of people, and the demands created by crop rotation for different varieties that will do well in different fields.

The potato demonstrates a well-known example of the risks of low crop diversity: during the tragic Irish potato famine (1845–1852 AD), the single potato variety grown in Ireland became susceptible to a potato blight—wiping out the crop (figure 4.1.3.*b*). The loss of the crop led to famine, death, and mass emigration. Resistance to disease is a chief benefit to maintaining crop biodiversity and lack of diversity in contemporary crop species carries similar risks. Seed companies, which are the source of most crop varieties in developed countries, must continually breed new varieties to keep up with evolving pest organisms. These same seed companies, however, have participated in the decline of the number of varieties available as they focus on selling fewer varieties in more areas of the world replacing traditional local varieties.



Figure 4.1.3.*b*: The effects of potato blight, which is cause by the water mold *Phytophthora infestans*. Image by USDA (public domain).

Potatoes are only one example of agricultural diversity. Every plant, animal, and fungus that has been cultivated by humans has been bred from original wild ancestor species into diverse varieties arising from the demands for food value, adaptation to growing conditions, and resistance to pests. The ability to create new crop varieties relies on the diversity of varieties available and the availability of wild forms related to the crop plant. These wild forms are often the source of new gene variants that can be bred with existing varieties to create varieties with new attributes. Loss of wild species related to a crop will mean the loss of potential in crop improvement. Maintaining the genetic diversity of wild species related to domesticated species ensures our continued supply of food.

Since the 1920s, government agriculture departments have maintained seed banks of crop varieties as a way to maintain crop diversity. This system has flaws because over time seed varieties are lost through accidents and there is no way to replace them. In 2008, the Svalbard Global seed Vault, located on Spitsbergen island, Norway, (figure 4.1.3.*c*) began storing seeds from around the world as a backup system to the regional seed banks. If a regional seed bank stores varieties in Svalbard, losses can be replaced from Svalbard should something happen to the regional seeds. The Svalbard seed vault is deep into the rock of the arctic island. Conditions within the vault are maintained at ideal temperature and humidity for seed survival, but the deep underground location of the vault in the arctic means that failure of the vault's systems will not compromise the climatic conditions inside the vault.



Figure 4.1.3.*c*: The Svalbard Global Seed Vault is a storage facility for seeds of Earth's diverse crops. (credit: Mari Tefre, Svalbard Global Seed Vault)

Supplemental Reading

Dunlap, Garrett. 2018. Facing Facts: Why a transmissible facial cancer is decimating Tasmanian devil populations. SITN Blog. The Graduate School of Arts and Sciences. Harvard University.

Attributions

Modified by Melissa Ha from the following sources:

- Biological from *AP Environmental Science* by University of California College Prep, University of California (licensed under CC-BY). Download for free at CNX.
- Importance of Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 4.1.3: Genetic Diversity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **9.3: Genetic Diversity** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

4.1.4: Patterns of Biodiversity

Biodiversity is not evenly distributed on the planet. For example, Lake Victoria in Africa (figure 4.1.4.*a*) contained almost 500 species from a family of fishes called cichlids before the introduction of the **invasive** Nile Perch in the 1980s and 1990s caused a mass extinction. (Note that this number does not include species of other fish families.) Lake Huron, the second largest of North America's Great Lakes, contains about 79 species of fish, all of which are found in many other lakes in North America. What accounts for the difference in diversity between Lake Victoria and Lake Huron? Lake Victoria is a tropical lake, while Lake Huron is a temperate lake. Lake Huron in its present form is only about 7,000 years old, while Lake Victoria in its present form is about 15,000 years old. These two factors, latitude and age, are two of several hypotheses biogeographers have suggested to explain biodiversity patterns on Earth.

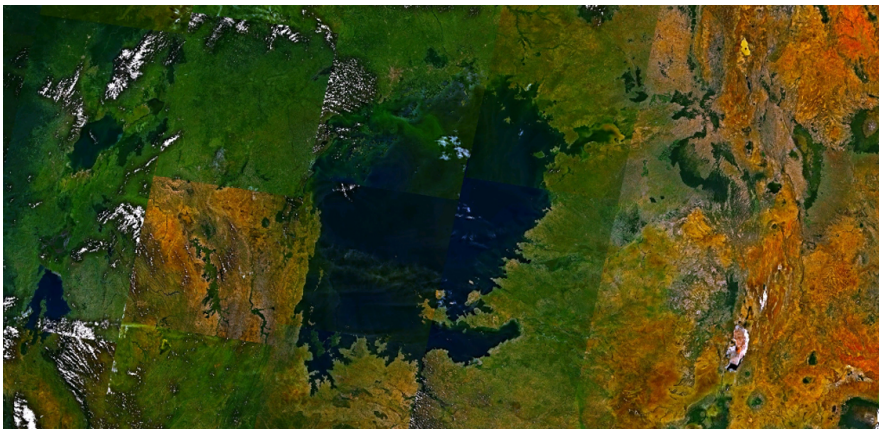


Figure 4.1.4.*a*: Lake Victoria in Africa, shown in this satellite image, was the site of one of the most extraordinary evolutionary findings on the planet, as well as a casualty of devastating biodiversity loss. (credit: modification of work by Rishabh Tatiraju, using NASA World Wind software)

Biogeography is the study of the distribution of the world's species both in the past and in the present. The work of biogeographers is critical to understanding our physical environment, how the environment affects species, and how changes in environment impact the distribution of a species. There are three main subfields of biogeography: ecological biogeography, historical biogeography (called paleobiogeography), and conservation biogeography. **Ecological biogeography** studies the current factors affecting the distribution of plants and animals. **Historical biogeography**, as the name implies, studies the past distribution of species. **Conservation biogeography**, on the other hand, is focused on the protection and restoration of species based upon the known historical and current ecological information.

The Tropics Have High Biodiversity

One of the oldest observed patterns in ecology is that biodiversity typically increases as latitude declines. In other words, the greatest species richness occurs near the equator and the lowest richness occurs near the poles (figure 4.1.4.*b*).

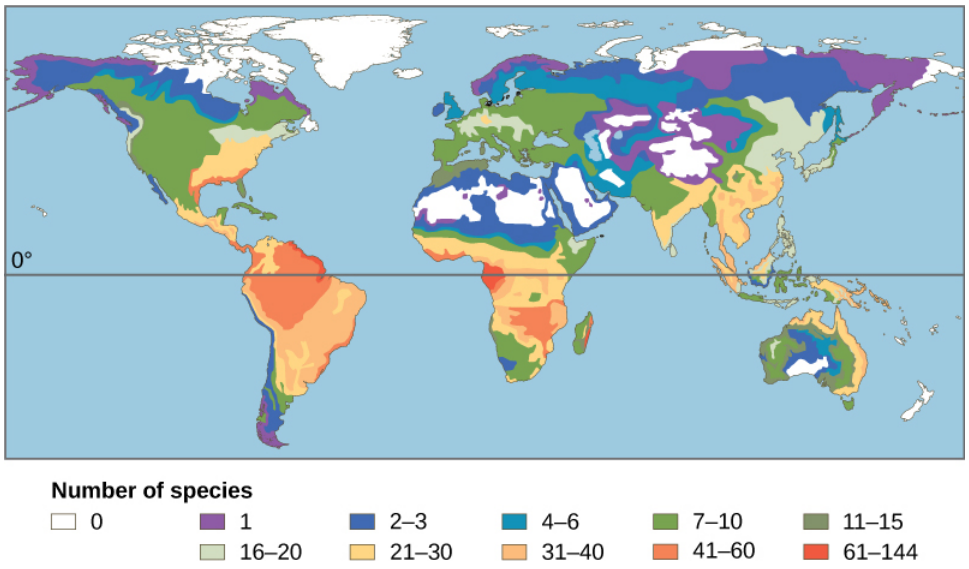


Figure 4.1.4.*b*: This map illustrates the number of amphibian species across the globe and shows the trend toward higher biodiversity at lower latitudes. A similar pattern is observed for most taxonomic groups. The greatest number of species, 61-144, are found in the Amazon region of South America and in parts of Africa. Between 21 and 60 species are found in other parts of South America and Africa, and in the eastern United States and Southeast Asia. Other parts of the world have between 1 and 20 amphibian species, with the fewest species occurring at northern and southern latitudes.

It is not yet clear why biodiversity increases closer to the equator, but hypotheses include the greater age of the ecosystems in the tropics versus temperate regions, which were largely devoid of life or drastically impoverished during the last ice age. The greater age provides more time for **speciation,** the evolutionary process of creating new species. Another possible explanation is the greater energy the tropics receive from the sun, but scientists have not been able to explain how greater energy input could translate into more species. The complexity of tropical ecosystems may promote speciation by increasing the habitat complexity, thus providing more **ecological niches** (see Communities chapter). Lastly, the tropics have been perceived as being more stable than temperate regions, which have a pronounced climate and day-length seasonality. The stability of tropical ecosystems might promote speciation. Regardless of the mechanisms, it is certainly true that biodiversity is greatest in the tropics. There are also high numbers of endemic species.

Biodiversity Hotspots

In 1988, British environmentalist Norman Myers developed a conservation concept to identify geographical areas rich in species and at significant risk for species loss: **biodiversity hotspots**. The original criteria for a hotspot included the presence of 1500 or more endemic plant species and 70 percent of the area disturbed by human activity. **Endemic species** are found in only one location. For example, all of cichlid species found in Lake Victoria only occurred there. The blue jay is endemic to North America, while the Barton Springs salamander is endemic to the mouth of one spring in Austin, Texas. Endemic species with highly restricted distributions, like the Barton Springs salamander, are particularly vulnerable to extinction. If a population of a widespread species declines in one region, individuals from another region may be able to recolonize the first location, but this is not possible for endemic species. Endemic species are particularly common in isolated regions, such as mountaintops or islands. Endemism is especially likely on islands that are large and far from the mainland. Identifying biodiversity hotspots aids with conservation efforts, a kind of conservation triage. By protecting hotspots, governments are able to protect a larger number of species. There are now 34 biodiversity hotspots (figure 4.1.4.*c*) containing large numbers of endemic species, which include half of Earth's endemic plants.

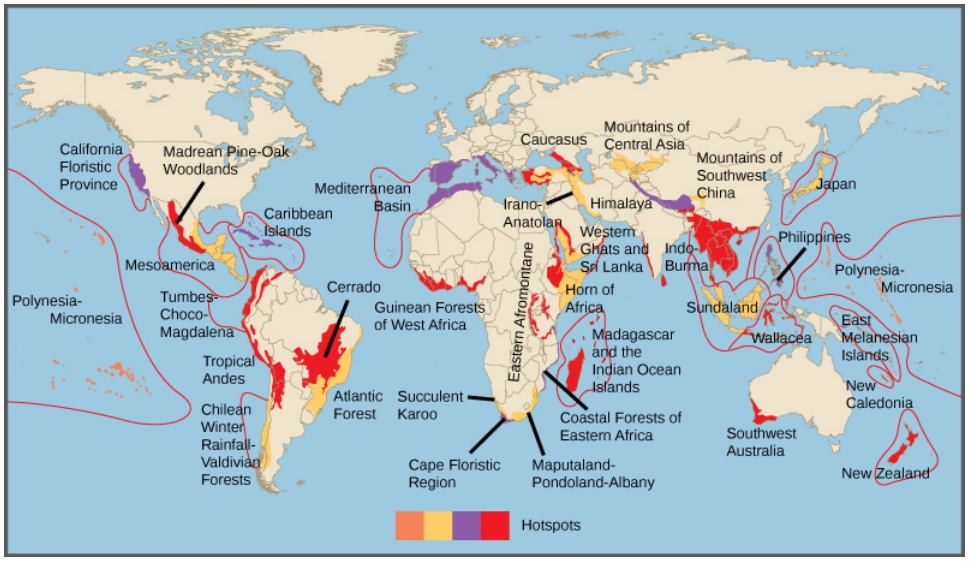


Figure 4.1.4.*c*: Conservation International has identified 34 biodiversity hotspots, which cover only 2.3 percent of the Earth's surface but have endemic to them 42 percent of the terrestrial vertebrate species and 50 percent of the world's plants. Hotspots are shaded in red, purple, yellow, or orange. Examples include the California Floristic Province, Tropical Andes, Mediterranean Basic, Cape Floristic Province, and Polynesia-Micronesia.

Biodiversity Hotspots in South Africa

A rich species diversity is found in South Africa. With a land surface of approximately 148,000 km^2 (about 57,000 square miles), representing approximately 1% of the Earth's total surface, South Africa contains 10% of the world's total known bird, fish and plant species, and 6% of the world's mammal and reptile species.

This natural wealth is threatened by the expansion of the human population and the increasing demand this places on the environment. The Karoo (figure 4.1.4.*d*) and the Cape are biodiversity hotspots in South Africa. South Africa has a wide range of climatic conditions and many variations in landscape, creating a variety of ecological niches, which can promote species diversity. These various landscapes give rise to the biomes which allow a wide variety of life to survive.



Figure 4.1.4.*d*: Map of South Africa, illustrating the Karoo, a dry region with hot summers and cold winters. Image by (WT-shared) NJR ZA at wts wikivoyage (CC-BY-SA).

More than 20,300 species of flowering plants occur in South Africa (figure 4.1.4.*e*). One of the six most important areas of plant growth in the world is in the Western Cape. Most of the 2,000 threatened species of plants are found in the **fynbos** in South-West Cape (figure 4.1.4.*f*).



Figure 4.1.4.*e*: Vegetation of the Karoo in near Willowmore, Eastern Cape, South Africa. Image by Winfried Bruenken (CC-BY-SA).



Figure 4.1.4.*f*: Vegetation of the fynbos in Western Cape, South Africa. Image by Joachim Huber (CC-BY-SA).

South Africa is home to 243 mammal species. Among the 17 threatened species in South Africa are the black rhino, pangolin, and giant golden mole. The blue antelope and quagga (figure 4.1.4.*g*) have become extinct. Of more than 800 bird species, 26 are threatened, including the penguin, Cape vulture, martial eagle (figure 4.1.4.*h*), and Cape parrot. In total 370 reptiles and amphibians occur in the region of which 21 are threatened. Six of these are endangered. Of the 220 freshwater fishes in South Africa, 21 are threatened. There are more than 2,000 marine fish species. While many insect species remain unidentified, 80,000 insects are known to occur.

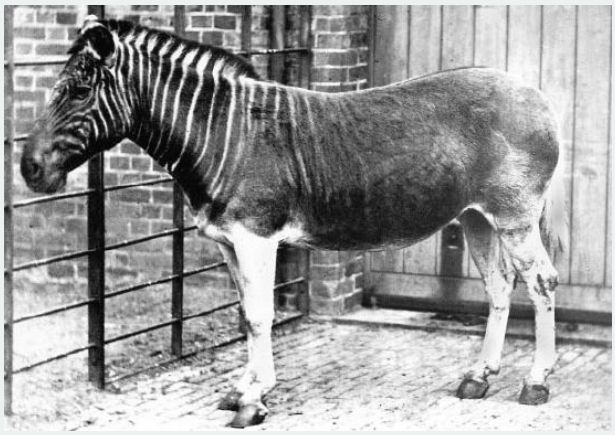


Figure 4.1.4.*g*: The quagga (*Equus quagga quagga*), now extinct, is technically a subspecies of the plains zebra. Image by Stanford Woods (public domain).



Figure 4.1.4.*h*: The Martial Eagle (*Polemaetus bellicosus*) is one of the threatened bird species found in South Africa. Image by Charles J. Sharp (CC-BY-SA).

Modified by Melissa Ha from Biodiversity and Biomes from *Life Sciences Grade 10* by Siyavula (CC-BY)

Spatial and Temporal Heterogeneity Promote Biodiversity

Environmental heterogeneity refers to the lack of uniformity in an area. Regions that are spatially heterogeneous are composed of smaller areas with contrasting vegetation, water and nutrient availability, soils and rocks, or other features. In regions that are temporally heterogeneous, conditions fluctuate over time. Heterogeneity promotes species richness by increasing the number of ecological niches, and heterogenous regions are thus priorities with respect to conserving biodiversity. Some regions are more hetergeneous than others. For example, a wetland consists of many microhabitats that differ in soil type and water availability, and water levels change over the seasons. Other regions, like the temperate grasslands in the midwestern United States, are relatively homogenous (uniform). Just as biodiversity can be assessed at a variety of scales, heterogeneity can be measured at large scales (such as within an entire biome) or small scales (such as within a single meadow).

Attributions

Modified by Melissa Ha from the following sources:

- Importance of Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- The Biodiverisity Crisis and Prelude to Conservation Biology and Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)

This page titled 4.1.4: Patterns of Biodiversity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **9.4: Patterns of Biodiversity** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

4.1.5: Data Dive- Biodiversity and Drugs

Overview

Global biodiversity continues to contribute significantly to the production of drugs. A 2016 paper reviewed drug sources for both synthetic and natural categories from 1981 to 2014. The figure below below displays the categories used in the analysis and a graph breaking down their results:

- “B”: Biological macromolecule
- “N”: Unaltered natural product, however, could have been semi- or totally synthetic
- “NB”: Natural product “botanical drug”
- “ND”: Natural product derivative
- “S”: Synthetic drug
- “S*”: Synthetic drug (NP pharmacophore)
- “V”: Vaccine
- “S-VM”: Natural product mimic of “S”
- “S*-VM”: Natural product mimic of “S*”

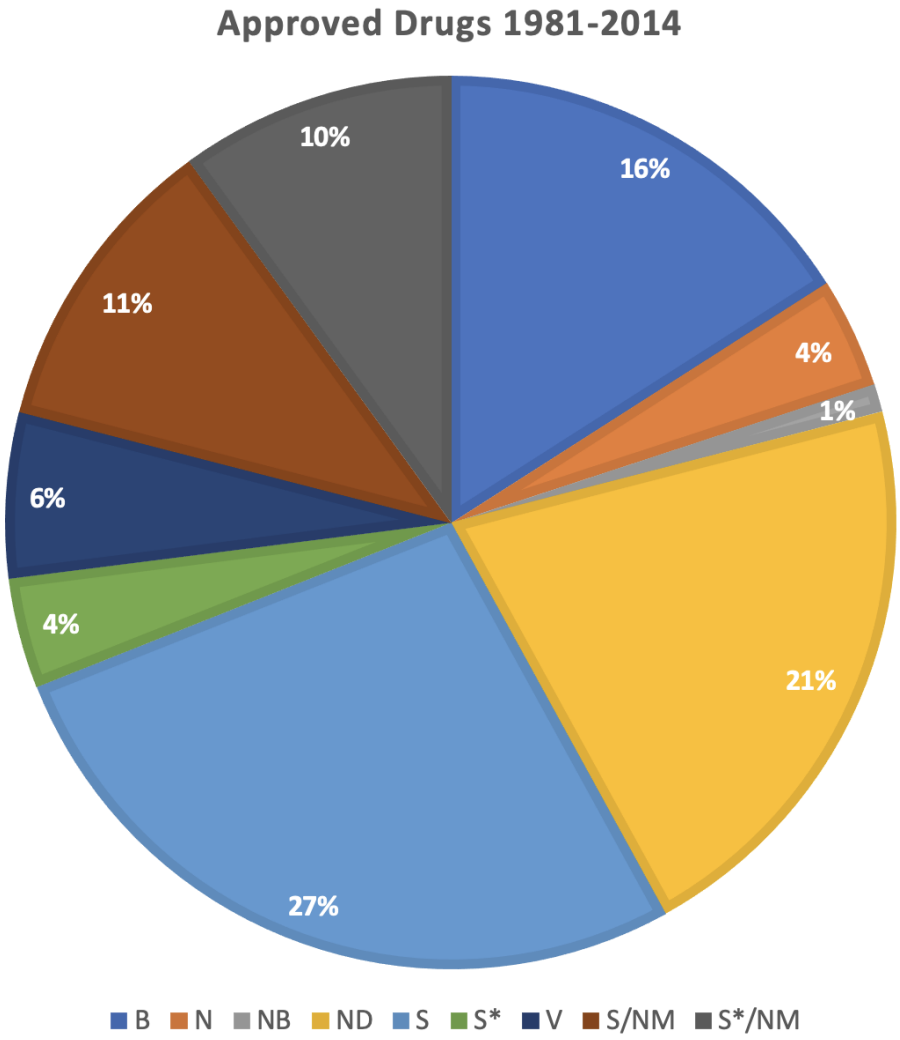


Figure 4.1.5.a: Percent of drugs approved between 1981 and 2004 with either a synthetic or natural sub-classification. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Newman DJ and Cragg GM 2016.

Questions

1. What question(s) are the authors trying to answer with this graph?
2. What are the top three largest approved drug categories?
3. After reading over the categories outlined, provide your impression/reflection for all the categories that were identified.
4. What reason could you provide as to why the N and NB categories are so small compared to ND?
5. Do you think the results of this graph can be used as a defense of saving/conserving biodiversity now and into the future? Why?

Raw Data From Above Graph(s)

Table 4.1.5.*a***:** Raw data for percent of drugs approved between 1981 and 2004 with either a synthetic or natural sub-classification. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Newman DJ and Cragg GM 2016.

This page titled 4.1.5: Data Dive- Biodiversity and Drugs is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **9.5: Data Dive- Biodiversity and Drugs** by Melissa Ha and Rachel Schleiger has no license indicated.

4.1.6: Review

Summary

After completing this chapter you should be able to...

- Define biodiversity.
- Distinguish among ecosystem, species, and genetic diversity, explaining the value of each.
- Define and provide examples of ecosystem services.
- Distinguish between species richness and species evenness.
- Provide an approximate estimate of the number of species on Earth as well as the percentage of species that have been identified.
- Explain the criteria for biodiversity hotspots and their role in conservation.
- Identify the characteristics of endemic species.
- Summarize hypotheses explaining high biodiversity near the equator.

Biodiversity exists at multiple levels of organization and is measured in different ways depending on the goals of those taking the measurements. These measurements include **ecosystem diversity**, **species diversity**, and **genetic diversity**. The number of described species is estimated to be 1.5 million with about 17,000 new species being described each year. Estimates for the total number of species on Earth vary but are on the order of 10 million. Biodiversity is negatively correlated with latitude for most taxa, meaning that biodiversity is higher in the tropics. The mechanism for this pattern is not known with certainty, but several plausible hypotheses have been advanced.

Attributions

Modified by Melissa Ha from Conservation & Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 4.1.6: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **9.6: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

4.2: Threats to Biodiversity

Chapter Hook

The northern spotted owl (*Strix occidentalis caurina*) is a western North American owl species that prefers large swaths of pristine old growth forests (that typically take 150-200 years to mature) for nesting. Unfortunately, most western forests have been regularly harvested for timber since around the establishment of the Forest Service in 1905. Thus, this species preferred habitat has been largely destroyed and their populations have plummeted. Habitat destruction is the number one cause for species extinction globally. For some species that are very habitat specific, like the northern spotted owl, there are not as many conservation actions to choose from that can help reverse declining population trends.



Figure 4.2.*a* Northern spotted owl perched in a tree. Image by USFS (licensed under CC-BY 2.0)

Biodiversity loss refers to the reduction of biodiversity due to displacement or extinction of species. The loss of a particular individual species may seem unimportant to some, especially if it is not a charismatic species like the Bengal tiger or the bottlenose dolphin. However, biologists estimate that species extinctions are currently many times higher the normal, or background, rate seen previously in Earth's history. This translates to the loss of tens of thousands of species within our lifetimes. This is likely to have

dramatic effects on human welfare through the collapse of ecosystems. Loss of biodiversity may have reverberating consequences on ecosystems because of the complex interrelations among species. For example, the extinction of one species may cause the extinction of another. To measure biodiversity loss, scientists assess which species are at risk of extinction as well as survey ecosystem decline.

The core threat to biodiversity on the planet is the combination of human population growth and the resources used by that population. The human population requires resources to survive and grow, and many of those resources are being removed unsustainably from the environment. The five main threats to biodiversity are habitat loss, pollution, overexploitation, invasive species, and climate change. Increased mobility and trade has resulted in the introduction of invasive species while the other threats are direct results of human population growth and resource use.

Attribution

Modified by Rachel Schleiger and Melissa Ha from Threats to Biodiversity and Importance of Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

Threats to Biodiversity

4.2.1: Extinction

4.2.2: Measures of Biodiversity Loss

4.2.3: Habitat Loss

4.2.4: Overexploitation

4.2.5: Pollution

4.2.6: Invasive Species

4.2.7: Climate Change

4.2.8: Data Dive- Threats to Biodiversity

4.2.9: Review

This page titled 4.2: Threats to Biodiversity is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

4.2.1: Extinction

Speciation and Extinction

The number of species on the planet, or in any geographical area, is the result of an equilibrium of two evolutionary processes that are ongoing: speciation and extinction. **Speciation** occurs when new species evolve, and **extinction** is the global loss of a species. When speciation rates begin to outstrip extinction rates, the number of species will increase; likewise, the number of species will decrease when extinction rates begin to overtake speciation rates. Throughout Earth's history, these two processes have fluctuated—sometimes leading to dramatic changes in the number of species on Earth as reflected in the fossil record (Figure 4.2.1.a).

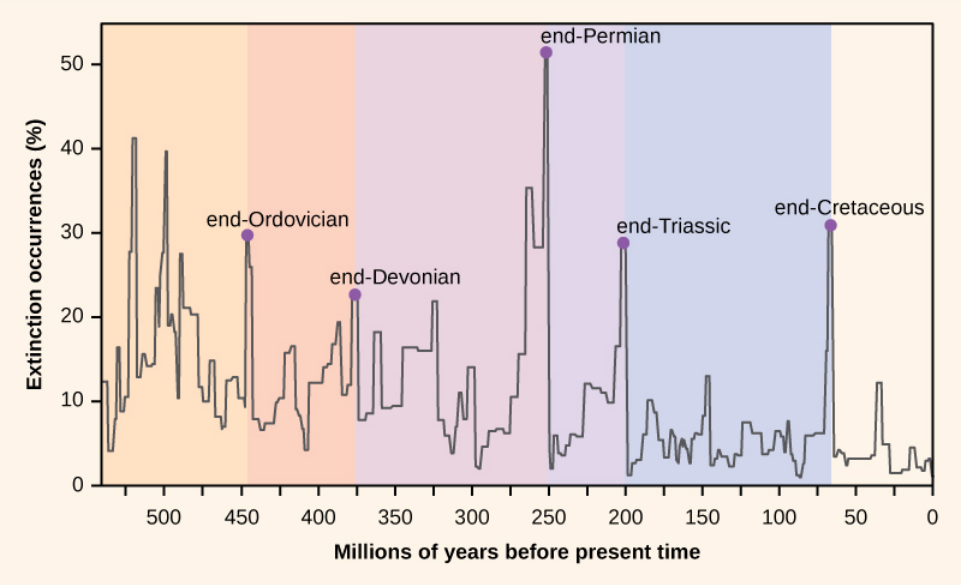


Figure 4.2.1.a: Percent extinction occurrences as reflected in the fossil record have fluctuated throughout Earth's history. This graph shows the extinction occurrences as a percentage over time (in millions of years before present time). At the lowest points on the cycle, extinction occurrences were between 2% and 5% percent. Sudden and dramatic losses of biodiversity, called mass extinctions, have occurred five times, marking the end of geological periods: end-Ordovician, 450 million years ago; end-Devonian, 374 million years ago; end-Permian, 252 million years ago; end-Triassic, 200 million years ago; and end-Cretaceous, 65 million years ago. During these spikes, extinction occurrences approximately ranged from 22% to 50%.

Mass Extinctions

Paleontologists have identified five events in geological history of sudden and dramatic losses in biodiversity, with more than half of all extant species disappearing from the fossil record. **Extant** species are those that are surviving (not extinct) at a point in time. These five events are called mass extinctions. There are many lesser, yet still dramatic, extinction events, but the five mass extinctions have attracted the most research. An argument can be made that the five mass extinctions are only the five most extreme events in a continuous series of large extinction events that have occurred since 542 million years ago.

The fossil record of the mass extinctions was the basis for defining periods of geological history, so they typically occur at the transition point between geological periods. The transition in fossils from one period to another reflects the dramatic loss of species and the gradual origin of new species. These transitions can be seen in the rock layers. Table 4.2.1.a provides the names and dates of the five mass extinctions.

In most cases, the hypothesized causes are still controversial; however, the causes of the most recent event, the end-Cretaceous extinction, are best understood (table 4.2.1.a). It was during this extinction event that the dinosaurs, the dominant vertebrate group for millions of years, disappeared from the planet (with the exception of a theropod clade that gave rise to birds). Indeed, every land animal that weighed more then 25 kg (55 lbs) became extinct. The cause of this extinction is now understood to be the result of a cataclysmic impact of a large meteorite, or asteroid, off the coast of what is now the Yucatán Peninsula. Biodiversity recovery times after mass extinctions vary, but have been up to 30 million years.

Table 4.2.1.*a*: Summary of the five mass extinctions, including the name, dates, percent of biodiversity lost, and hypothesized causes.

Geological Period	Mass Extinction Name	Time (millions of years ago)	Loss in Biodiversity	Hypothesized Cause(s)
Ordovician–Silurian	end-Ordovician O–S	450–440	85% of marine species	Global cooling and then warming, gamma-ray burst from supernova removed ozone layer
Late Devonian	end-Devonian	375–360	79-87% of all species	Unknown
Permian–Triassic	end-Permian	251	96% of marine species and 70% of terrestrial (land) species	Volcanic activity, decrease on dissolved oxygen in the oceans
Triassic–Jurassic	end-Triassic	205	76% of all species	Climate change, asteroid impact, volcanic eruptions
Cretaceous–Paleogene	end-Cretaceous K–Pg (K–T)	65.5	50% of animals and plants	Asteroid impact

The Pleistocene Extinction is one of the lesser extinctions, and a recent one. It is well known that the North American, and to some degree Eurasian **megafauna**, or large animals, disappeared toward the end of the last glaciation (cooling) period. The extinction appears to have happened in a relatively restricted time period of 10,000–12,000 years ago. In North America, the losses were quite dramatic and included the woolly mammoths (last dated about 4,000 years ago in an isolated population), mastodon, giant beavers, giant ground sloths, saber-toothed cats, and the North American camel, just to name a few. The possibility that the rapid extinction of these large animals was caused by overhunting, a type of overexploitation, was first suggested in the 1900s. Research into this hypothesis continues today. It seems likely that overhunting caused many pre-written history extinctions in many regions of the world.

The Sixth Mass Extinction

The sixth, or Holocene, mass extinction appears to have begun earlier than previously believed and has mostly to do with the activities of *Homo sapiens*. Since the beginning of the Holocene period, there are numerous recent extinctions of individual species that are recorded in human writings. Most of these are coincident with the expansion of the European colonies since the 1500s.

One of the earlier and popularly known examples is the dodo bird. The dodo bird lived in the forests of Mauritius, an island in the Indian Ocean. The dodo bird became extinct around 1662 (figure 4.2.1.*b*). It was hunted for its meat by sailors and was easy prey because the dodo, which did not evolve with humans, would approach people without fear. Introduced pigs, rats, and dogs brought to the island by European ships also killed dodo young and eggs.



Figure 4.2.1.*b*: The dodo bird was hunted to extinction around 1662. (credit: Ed Uthman, taken in Natural History Museum, London, England)

Steller's sea cow became extinct in 1768; it was related to the manatee and probably once lived along the northwest coast of North America. Steller's sea cow was first discovered by Europeans in 1741 and was hunted for meat and oil. The last sea cow was killed in 1768. That amounts to 27 years between the sea cow's first contact with Europeans and extinction of the species.

In 1914, the last living passenger pigeon died in a zoo in Cincinnati, Ohio. This species had once darkened the skies of North America during its migrations, but it was hunted and suffered from habitat loss through the clearing of forests for farmland. In 1918, the last living Carolina parakeet died in captivity. This species was once common in the eastern United States, but it suffered from habitat loss. The species was also hunted because it ate orchard fruit when its native foods were destroyed to make way for farmland. The Japanese sea lion, which inhabited a broad area around Japan and the coast of Korea, became extinct in the 1950s due to fishermen. The Caribbean monk seal was distributed throughout the Caribbean Sea but was driven to extinction via hunting by 1952.

As described in the next section, the current high rates of extinction will cause a large and rapid decline in the biodiversity of the planet. According to a 2019 United Nations report, 1 million species are at risk of extinction. Considering there are estimated to be 8-11 million species total (see The Number of Species on Earth), that means up to 12.5% of species could go extinct, and many of them within our lifetimes.

Estimates of Present-Time Extinction Rates

The **background extinction rate** is estimated to be about 1 extinction per million species each year (E/MSY). For example, if there are 8-11 million species in existence, then we would expect 8-11 of those species to become extinct in a year.

Estimates of extinction rates are hampered by the fact that most extinctions are probably happening without observation. One contemporary extinction-rate estimate uses the extinctions in the written record since the year 1500. For birds alone, this method yields an estimate of 26 E/MSY, almost three times the background rate. However, this value may be underestimated for three reasons. First, many existing species would not have been described until much later in the time period and so their loss would have gone unnoticed. Second, we know the number is higher than the written record suggests because now extinct species are being described from skeletal remains that were never mentioned in written history. And third, some species are probably already extinct even though conservationists are reluctant to name them as such. Taking these factors into account raises the estimated extinction rate to nearer 100 E/MSY. The predicted rate by the end of the century is 1500 E/MSY.

A second approach to estimating present-time extinction rates is to correlate species loss with habitat loss, and it is based on measuring forest-area loss and understanding species–area relationships. The species-area relationship is the rate at which new species are seen when the area surveyed is increased (figure 4.2.1.*c*). Likewise, if the habitat area is reduced, the number of species

seen will also decline. This kind of relationship is also seen in the relationship between an island's area and the number of species present on the island: as one increases, so does the other, though not in a straight line. Estimates of extinction rates based on habitat loss and species–area relationships have suggested that with about 90 percent of habitat loss an expected 50 percent of species would become extinct. Species–area estimates have led to estimates of present-day species extinction rates of about 1000 E/MSY and higher.

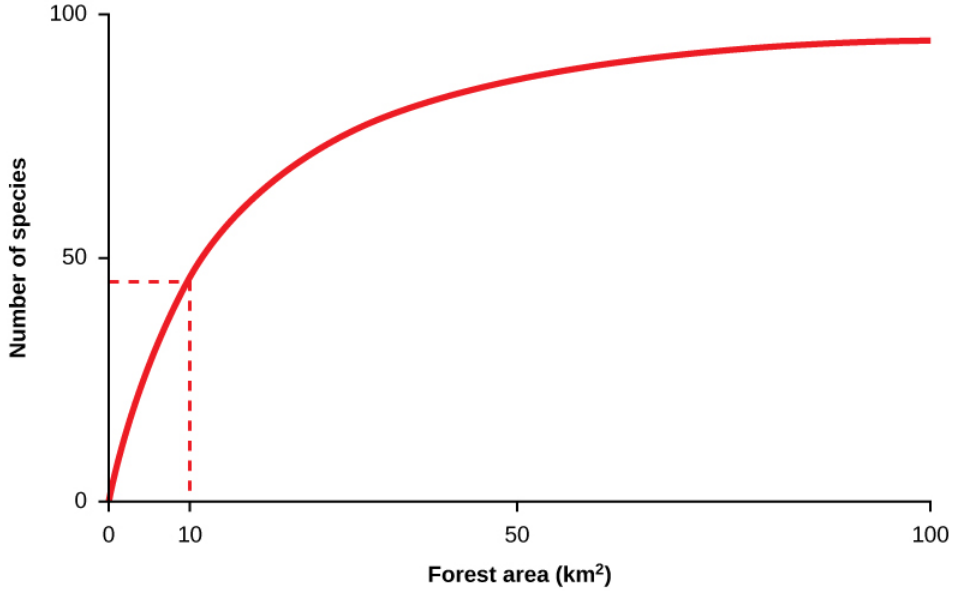


Figure 4.2.1.*c*: A typical species-area curve shows the cumulative number of species found as more and more area is sampled. The curve has also been interpreted to show the effect on species numbers of destroying habitat; a reduction in habitat of 90 percent from 100 km^2 to 10 km^2 reduces the number of species supported by about 50 percent.

Many extinctions will affect species that biologists have not yet discovered. Most of these "invisible" species that will become extinct currently live in tropical rainforests (figure 4.2.1.*d*). These rainforests are the most diverse ecosystems on the planet and are being destroyed rapidly by deforestation to provide timber and space for agriculture.



Figure 4.2.1.*d*: This tropical lowland rainforest in Madagascar is an example of a high biodiversity habitat. This particular location is protected within a national forest, yet only 10 percent of the original coastal lowland forest remains, and research suggests half the original biodiversity has been lost. (credit: Frank Vassen)

Extirpation

The elimination of species at a local level–known as **extirpation** – also poses threats to the integrity and sustainability of ecosystems. Widespread extirpation can obviously lead to extinction, but absence of species, even at a local scale, can affect ecosystem function. For example, by the mid-1920s wolves had been extirpated from Yellowstone National Park, although they continued to thrive elsewhere. When wolves were reintroduced to the park in the mid-1990s, they regulated elk populations,

benefiting the vegetation and plant communities (see Ecosystem Restoration). What mattered for ecosystem function in Yellowstone was whether wolves were present there, not just whether the species survived somewhere.

References

UN Report: Nature's Dangerous Decline 'Unprecedented'; Species Extinction Rates 'Accelerating'. 2019. United Nations. Accessed 2020-08-01.

Attributions

Modified by Melissa Ha from the following sources:

- The Biodiversity Crisis from *General Biology* by OpenStax (licensed under CC-BY)
- Preserving Biodiversity and Importance of Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Biodiversity, Species Loss, and Ecosystem Function by Tom Theis and Jonathan Tomkin, Editors. Download for free at CNX. (licensed under CC-BY)

This page titled 4.2.1: Extinction is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **10.1: Extinction** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

4.2.2: Measures of Biodiversity Loss

The Red List

The **International Union for the Conservation of Nature (IUCN)** coordinates efforts to catalog and preserve biodiversity worldwide. One way scientists gauge trends in biodiversity is by monitoring the fate of individual species. Since 1964, the IUCN has compiled information in the Red List of Threatened Species, which includes plants, animals, fungi, and selected brown algae species. Updates to the Red List are released every four years.

Species can be classified into nine Red List categories based on their extinction risk. Firstly, there are species that are already **extinct** and those that are **extinct in the wild**, meaning that the remaining individuals are only found in captivity. Species at risk of extinction are called **threatened**. Species that are at risk of becoming threatened are called **near threatened**. The Emperor Penguin (*Aptenodytes forsteri*) is an example a near threatened species, mainly due to habitat loss and climate change (figure 4.2.2.a). Those with little risk of extinction are designated **least concern**. Note that only a fraction of the 8-11 million species on Earth are identified (see The Number of Species on Earth). For many of the species that are identified, data still needs to be gathered before they can be assigned to a Red List category (**data deficient**). Less than 10% of the approximately 1.5 million identified species have been assessed for the Red List at all. Species that have not been assessed are deemed **not evaluated**.



Figure 4.2.2.a: The Emperor Penguin is an example of near threatened species. Image by Christopher Michel (CC-BY).

There are three categories of threatened species: **vulnerable, endangered,** and **critically endangered**. Of these, critically endangered species have the greatest risk while vulnerable species are at the least risk out of the threatened categories. The African elephant (*Loxodonta africana*) is listed vulnerable species due to poaching and habitat loss (figure 4.2.2.b). Interestingly, the forest subspecies (*L. africana cylotis*) is at much greater risk than the savanna subspecies (*L. africana africana*), but the Red List category is assigned at the species level in this case. The blue whale (*Balaenoptera musculus*) is endangered due to pollution, climate change, and poaching (figure 4.2.2.c). Fortunately, blue whale population size is increasing. Several species of pitcher plants, which trap insects as a source of nitrogen, are critically endangered due to overexploitation and habitat loss (figure 4.2.2.d).



Figure 4.2.2.*b*: The African forest elephant (*Loxodonta africana cylotis*) is at greater risk of extinction than the savanna subspecies. When both subspecies are grouped together, African elephants are considered vulnerable. Image by Peter H. Wrege (CC-BY-SA).



Figure 4.2.2.*c*: The blue whale (*Balaenoptera musculus*) is categorized as endangered according to the IUCN Red List. Image by Mike Baird (CC-BY).



Figure 4.2.2.*d*: *Nepenthes aristolochioides* is a critically endangered pitcher plant. Image by Alfindra Primaldhi (CC-BY).

Extinct and Threatened Animals

Scientists know much more about the state of **vertebrates**—especially mammals, birds, and amphibians—than they do about other forms of animal life. Of 6,594 described mammal species, 96 species have become extinct during the last 500 years (Mammal Diversity Database). According to the Red List, approximately 26% of mammal species worldwide are known to be threatened. A smaller percentage (about 14%) of the world's 10,721 identified bird species (Birds of the World) are threatened (figure 4.2.2.*e*).



Figure 4.2.2.*e*: North American passenger pigeons lived in enormous flocks and were once the most numerous birds on earth. Market hunting on a massive scale and habitat destruction combined to extinguish them as a species in the early twentieth century. Image by James St. John (CC-BY).

Among the well-studied vertebrates, amphibians are faring especially poorly. Of the more than 6,000 known species of amphibians, 35 have become extinct worldwide since 1500 (figure 4.2.2.*f*), and two others are extinct in the wild (Red List). Overall, 41% of the world's amphibian species are known to be threatened (Red List). Only small proportions of the world's species of reptiles and fish have been evaluated for purposes of the Red List. Among those, 34% of selected reptiles and 8% of selected bony fishes are threatened (Red List).



Figure 4.2.2.*f*: The extinct golden toad of Monteverde, Costa Rica, was last seen in 1989. Its extinction is attributed to fungal infection combined with especially hot, dry weather. Source: U.S. Fish and Wildlife Service via Wikimedia Commons

Invertebrates comprise the vast majority of animals, an estimated 97% of animal species. They include everything from insects and arachnids, to mollusks, crustaceans, corals, and more. Few of these groups have been assessed in a comprehensive way, but assessments within some groups call attention to disturbing, large-scale trends. For example, 33% of the world's reef-building corals are already considered threatened (Red List), and many more of them are experiencing rates of decline that move them toward threatened status (figure 4.2.2.*g*). The demise of reef-building corals has magnified ecological impacts, since so much other marine life depends on them.

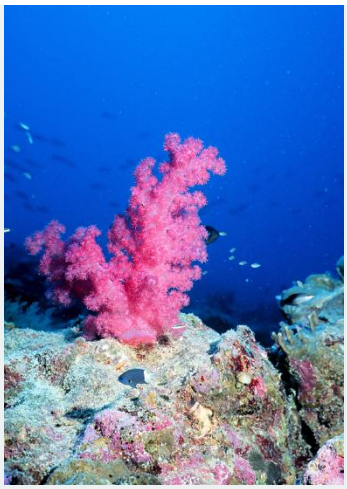


Figure 4.2.2.*g*: Reef-formal corals, such this pink soft coral, are at high risk of extinction. Source: Linda Wade via National Oceanic & Atmospheric Administration (NOAA)

Ecosystem Loss and Alteration

Another way of gauging biodiversity involves assessment on the scale of ecosystems. The causes of wholesale losses of ecosystems are much the same as those driving extinction or endangerment of species, with habitat destruction and fragmentation being the primary agent. Worldwide, for example, the conversion of land to agriculture and cultivation have led to significant losses in grassland ecosystems. In North America, nearly 70% of the tallgrass prairie ecosystem (which once covered 142 million acres) has been converted to agriculture, and losses from other causes, such as urban development, have brought the total to about 90%. Current estimates indicate that agricultural activity and cultivation systems now cover nearly 25% of the Earth's surface

According to the United Nations Millennium Ecosystem Assessment, by the beginning of the 21st century, 15 of the world's 24 ecosystems, from rainforests to aquifers to fisheries, were rated in serious decline. For example, tropical rainforests, which are the habitats for nearly half of the world's plant and animal species, covered about 4 billion acres in past centuries, but only 2.5 billion acres remain and nearly 1% is being lost annually. Losses have been especially severe in the tropics of Africa and Southeast Asia. Current estimates indicate that about 50% of the world's wetland habitat has been lost. (Note that wetlands are a broad collection of many different ecosytem types.) The former extent of wetland habitats worldwide (fresh, brackish and salt) is difficult to determine but certainly exceeded a billion acres.

References

Birds of the World. 2020. The Cornell Lab of Ornithology. Accessed 2020-07-29.

The IUCN Red List of Threatened Species. 2020. IUCN. Accessed 2020-07-29.

Mammal Diversity Database. 2020. American Society of Mammalogists. Accessed 2020-07-29.

Attribution

Modified by Melissa Ha from Biodiversity, Species Loss, and Ecosystem Function and The Industrialization of Nature: A Modern History (1500 to the present) from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors. Download for free at CNX. (licensed under CC-BY)

This page titled 4.2.2: Measures of Biodiversity Loss is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **10.2: Measures of Biodiversity Loss** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

4.2.3: Habitat Loss

Humans rely on technology to modify their environment and make it habitable. Other species cannot do this. Elimination of their habitat—whether it is a forest, coral reef, grassland, or flowing river—will kill the individuals in the species. Remove the entire habitat, and the species will become extinct, unless they are among the few species that do well in human-built environments. Habitat loss includes habitat destruction and habitat fragmentation.

Habitat Destruction

Habitat destruction occurs when the physical environment required by a species is altered so that the species can no longer live there. Human destruction of habitats accelerated in the latter half of the twentieth century. Consider the exceptional biodiversity of Sumatra: it is home to one species of orangutan, a species of critically endangered elephant, and the Sumatran tiger, but half of Sumatra's forest is now gone. The neighboring island of Borneo, home to the other species of orangutan, has lost a similar area of forest. Forest loss continues in protected areas of Borneo. The orangutan in Borneo is listed as endangered by the International Union for Conservation of Nature (IUCN), but it is simply the most visible of thousands of species that will not survive the disappearance of the forests of Borneo. The forests are removed for timber and to plant palm oil plantations (figure $4.2.3.a$). Palm oil is used in many products including food products, cosmetics, and biodiesel in Europe. According to Global Forest Watch, 9.7% of tree cover was lost globally from 2002 to 2019, and 9% of that occurred in Indonesia and Malaysia (where Sumatra and Borneo are located). Figure $4.2.3.b$ shows the average annual change in forest area around the world from 1990 to 2015.



(a)



(b)



(c)



(d)



(e)

Figure 4.2.3.a: (a) One species of orangutan, *Pongo pygmaeus*, is found only in the rainforests of Borneo, and the other species of orangutan (*Pongo abelii*) is found only in the rainforests of Sumatra. These animals are examples of the exceptional biodiversity of (c) the islands of Sumatra and Borneo, which are located in the south Pacific, just northwest of Australia. Sumatra is in the country of Indonesia. Half of Borneo is in Indonesia, and half is in Malaysia. Other species include the (b) Sumatran tiger (*Panthera tigris sumatrae*) and the (d) Sumatran elephant (*Elephas maximus sumatranus*), both critically endangered species. Rainforest habitat is being removed to make way for (e) oil palm plantations such as this one in Borneo's Sabah Province. (credit a: modification of work by Thorsten Bachner; credit b: modification of work by Dick Mudde; credit c: modification of work by U.S. CIA World Factbook; credit d: modification of work by "Nonprofit Organizations"/Flickr; credit e: modification of work by Dr. Lian Pin Koh)

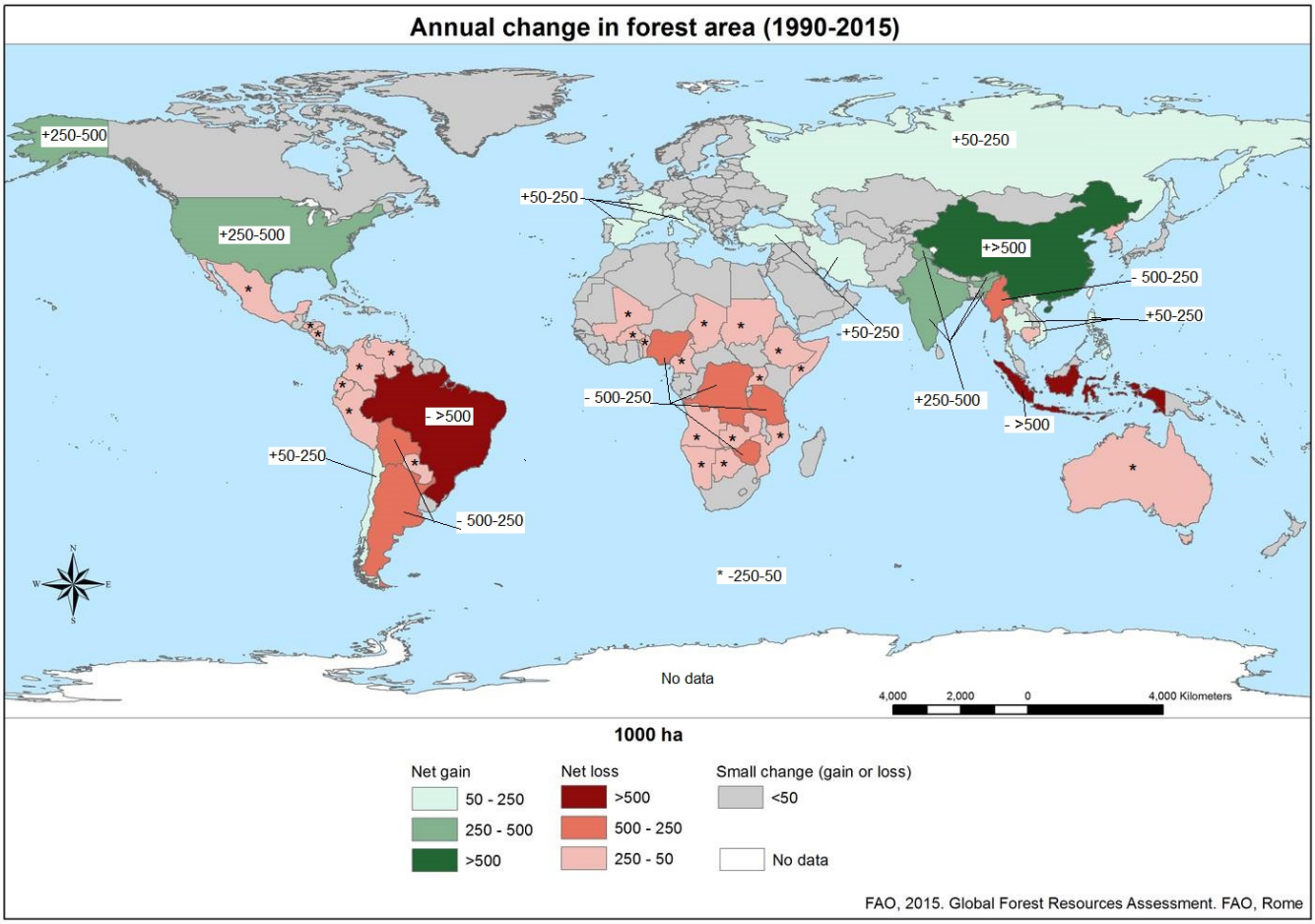


Figure 4.2.3.*b*: Average annual change in forest area globally from 2005 to 2015. China is dark green, indicating that have gained more than 500 "kilohectares" (kha) of forest. Medium green indicates countries that have gained 250-500 kha, including the the United States, India, Ghutan, and Bangladesh. Light green indicates countries that have gained 50-250 khA, including Chile, Spain, France, Italy, Turkey, Iran, Russia, Vietnam, and Thailand. Indonesia and Brazil lost more than 500 kha of forest (marked with dark red). Countries that lost 500-250 kha are marked with medium red, including Bolivia, Argentina, Nigeria, Democratic Republic of the Congo, Tanzania, Zimbabwe, and Myanmar. Countries that lost 250-50 kha are marked with light red and an asterisk (*). These include Mexico, Honduras, Nicaragua, Venezuela, Colombia, Ecuador, Peru, Paraguay, Mali, Burkina Faso, Benin, Cameroon, Chad, Sudan, Ethiopia, Somalia, Uganda, Angola, Namibia, Botswana, Zimbabwe, Mozambique, Cambodia, North Korea, and Australia. No data has been collected for Antarctica. All other regions have had lost or gained less than 50 kha in forest area. Image by FAO, 2015. Global Forest Resources Assessment. FAO. Rome (Open Access Policy).

Preventing Habitat Destruction with Wise Wood Choices

Most consumers do not imagine that the home improvement products they buy might be contributing to habitat loss and species extinctions. Yet the market for illegally harvested tropical timber is huge, and the wood products often find themselves in building supply stores in the United States. One estimate is that up to 10% of the imported timber in the United States, which is the world's largest consumer of wood products, is illegally logged. A 2012 United Nations and Interpol report estimated that the illegal timber trade is worth $30-100 billion each year. Most of the illegal products are imported from countries that act as intermediaries and are not the originators of the wood.

How is it possible to determine if a wood product, such as flooring, was harvested sustainably or even legally? The Forest Stewardship Council (FSC) certifies sustainably harvested forest products (figure 4.2.3.*c*). Looking for their certification on flooring and other hardwood products is one way to ensure that the wood has not been taken illegally from a tropical forest. There are certifications other than the FSC, but these are run by timber companies, thus creating a conflict of interest. Another approach is to buy domestic wood species. While it would be great if there was a list of legal versus illegal woods, it is not that simple. Logging and forest management laws vary from country to country; what is illegal in one country may be legal in another. Where and how a product is harvested and whether the forest from which it comes is being sustainably maintained all

factor into whether a wood product will be certified by the FSC. It is always a good idea to ask questions about where a wood product came from and how the supplier knows that it was harvested legally.



Figure 4.2.3.*c*: The sawmill in Uaxactun, Guatamala is Forest Stewardship Council (FSC) certified and provides good income from a sustainable resource for not only the saw operators but also many others who help keep the operation running. Image by Jason Houston for USAID (public domain).

Habitat Destruction of Rivers and Streams

Habitat destruction can affect ecosystems other than forests. Rivers and streams are important ecosystems and are frequently the target of habitat modification. Damming of rivers affects flow and access to habitat. Altering a flow regime can reduce or eliminate populations that are adapted to seasonal changes in flow. For example, an estimated 91% of riverways in the United States have been modified with damming or stream bank modification. Many fish species in the United States, especially rare species or species with restricted distributions, have seen declines caused by river damming and habitat loss. Research has confirmed that species of amphibians that must carry out parts of their life cycles in both aquatic and terrestrial habitats are at greater risk of population declines and extinction because of the increased likelihood that one of their habitats or access between them will be lost. This is of particular concern because amphibians have been declining in numbers and going extinct more rapidly than many other groups for a variety of possible reasons.

Habitat Fragmentation

Habitat fragmentation occurs when the living space of a species is divided into discontinuous patches. For example, a mountain highway could divide a forest habitat into separate patches. This is especially problematic for consumers at the top of the food chain, which require large ranges to find adequate prey. Additionally, habitat fragmentation separates individuals from potential mates. **Wildlife corridors** mitigate the damage of habitat fragmentation by connecting patches with suitable habitat. For example, the bridge over a highway could allow animals to move between habitat patches (figure 4.2.3.*d*). Riparian areas, areas of land adjacent to bodies of water, such as streams, can serve as natural wildlife corridors when left intact.



Figure 4.2.3.*d*: This overpass on the Trans-Canada Highway between Banff and Lake Louise, Alberta, serves as a wildlife corridor. Image by WikiPedant at Wikimedia Commons (CC-BY-SA).

References

Global Forest Watch. 2020. World Resources Institute. Accessed 2020-07-29.

Illegal logging nets organized crime up to 100 billion dollars a year, INTERPOL–UNEP report reveals. 2012. Interpol. Accessed 2020-07-29.

Attributions

Modified by Melissa Ha from the following sources:

- Threats to Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Threats to Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)

This page titled 4.2.3: Habitat Loss is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **10.3: Habitat Loss** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

4.2.4: Overexploitation

Overexploitation (overharvesting) involves hunting, fishing, or otherwise collecting organisms at a faster rate than they can be replenished.

Terrestrial Animals

Terrestrial animals may be overexploited as sources of food, garments, jewelry, medicine, or pets. For example, the poaching of elephants for their valuable ivory and rhinos for their horns, which are used in traditional medicine, is a major threat to these species. There are also concerns about the effect of the pet trade on some terrestrial species such as turtles, amphibians, birds, plants, and even the orangutans. Harvesting of pangolins for their scales and meat, and as curiosities, has led to a drastic decline in population size (figure 4.2.4.*a*).



Figure 4.2.4.*a*: Pangolins are threatened by overexploitation. This work by David Brossard is licensed under CC-BY.

Bush meat is the generic term used for wild animals killed for food. Hunting is practiced throughout the world, but hunting practices, particularly in equatorial Africa and parts of Asia, are believed to threaten several species with extinction. Traditionally, bush meat in Africa was hunted to feed families directly. However, recent commercialization of the practice now has bush meat available in grocery stores, which has increased harvest rates to the level of unsustainability. Additionally, human population growth has increased the need for protein foods that are not being met from agriculture. Species threatened by the bush meat trade are mostly mammals including many monkeys and the great apes living in the Congo basin.

Aquatic Animals

Aquatic species are particularly vulnerable to overexploitation, which is more specifically called **overfishing** in this case. For about one billion people, aquatic resources provide the main source of animal protein (figure 4.2.4.*b*), but since 1990, production from global **fisheries** (areas for catching wild or farmed fish or other aquatic animals) has declined. Figure 4.2.4.*c* illustrates the extent of overfishing in the U.S. Despite considerable effort, few fisheries are managed sustainability. For example, the western Atlantic cod fishery was a hugely productive fishery for 400 years, but the introduction of modern fishing vessels in the 1980s and the pressure on the fishery led to it becoming unsustainable. Bluefin tuna are in danger of extinction. The once-abundant Mediterranean swordfish fishery have been depleted to commercial and biological exhaustion.



Figure 4.2.4.*b*: Fishing boats at marine fishers. Alaskan waters have been fished by people for thousands of years, but they are under pressure from modern fishing technologies and large-scale extraction. Source: National Oceanic and Atmospheric Administration

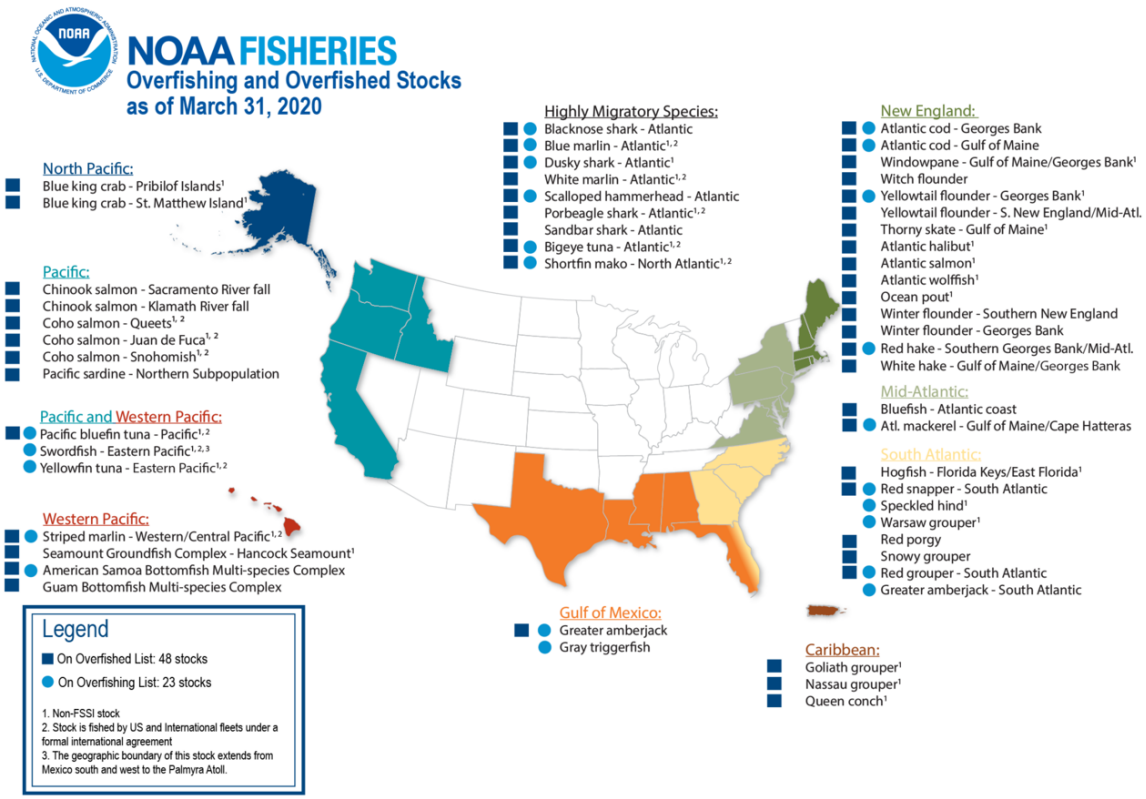


Figure 4.2.4.*c*: Map of overfishing and overfished stocks in the U.S. by region. Stocks on the overfishing list are being harvested too quickly, and those on the overfished list have population sizes that are too low. For example, stocks of Chinook salmon, Coho salmon, and Pacific sardines are overfished in the Pacific. Some species, including stocks of Pacific bluefin tuna and Atlantic cod, are on both the overfishing and overfished lists. Image by NOAA (public domain).

Most fisheries are managed as a common resource, available to anyone willing to fish, even when the fishing territory lies within a country's territorial waters. Common resources are subject to an economic pressure known as the **tragedy of the commons**, in which fishers have little motivation to exercise restraint in harvesting a fishery when they do not own the fishery. This results on

overexploitation. In a few fisheries, the biological growth of the resource is less than the potential growth of the profits made from fishing if that time and money were invested elsewhere. In these cases—whales are an example—economic forces will drive toward fishing the population to extinction.

Overfishing can result in a radical restructuring of the marine ecosystem in which a dominant species is so overexploited that it no longer serves its ecological function. For example, overfishing a tertiary consumer could causes populations of secondary consumers to increase. Secondary consumers would then feed on primary consumes (like zooplankton), decreasing their population size. With fewer zooplankton, populations of primary producers (**phytoplankton**, or photosynthetic microorganisms) would be unregulated (see Food Chains). The collapse of fisheries has dramatic and long-lasting effects on local human populations that work in the fishery. In addition, the loss of an inexpensive protein source to populations that cannot afford to replace it will increase the cost of living and limit societies in other ways. In general, the fish taken from fisheries have shifted to smaller species, and the larger species are overfished. The ultimate outcome could clearly be the loss of aquatic systems as food sources.

A related consequence of fishing practices is "**bycatch**," animals that fishers sometimes catch and discard because they do not want them, cannot sell them, or are not allowed to keep them. Bycatch can be fish, but also includes other animals such as dolphins, whales, sea turtles, and seabirds that become hooked or entangled in fishing gear.

Coral reefs are extremely diverse marine ecosystems that face peril from several processes. Reefs are home to 1/3 of the world's marine fish species—about 4,000 species—despite making up only one percent of marine habitat. Most home marine aquaria house coral reef species that are wild-caught organisms—not cultured organisms. Although no marine species is known to have been driven extinct by the pet trade, there are studies showing that populations of some species have declined in response to harvesting, indicating that the harvest is not sustainable at those levels.

Plants and Fungi

Some plant and fungal species are also overexploited, particularly if they are slow-growing. For example, stocks of wild ginseng, which is valued for its health benefits, are dwindling. Peyote cactus, which causes hallucinations and is used in sacred ceremonies, is also declining. Yarsagumba, dead moth larvae that were infected by fungal parasites (caterpillar fungus, *Ophiocordyceps sinensis*), is overexploited because it is highly valued in traditional medicine and used as an aphrodisiac (figure 4.2.4.*d*).



Figure 4.2.4.*d*: Yarsagumba is a combination of moth larvae and the fungus that infected and killed it. Image by Punya (CC-BY-SA).

Attributions

Modified by Melissa Ha from the following sources:

- Threats to Biodiversity and Importance of Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Threats to Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)
- Case Study: Marine Fisheries from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors. Download for free at CNX. (licensed under CC-BY)
- Understanding Bycatch. NOAA Fisheries. Accessed 28 March 2021 (public domain)

This page titled 4.2.4: Overexploitation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **10.4: Overexploitation** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

4.2.5: Pollution

Pollution occurs when chemicals, particles, or other materials are released into the environment, harming the organisms there. For many thousands of years, ever since they built the first campfire, human activity has generated air, water, and soil pollution. For most of human history, however, these contaminants had relatively little environmental impact. But over the last few centuries, pollution levels skyrocketed due to population growth and the Industrial Revolution. As a result, regulations have been enacted to control emissions. Even where these are effective in curbing current pollution sources, high levels of contamination may exist from past activity. And new contamination can occur through industrial accidents or other inadvertent releases of toxic substances.

Pollution has contributed to the decline of many threatened species. For example, a 2007 study by Kingsford and colleagues found that pollution was a major pressure on 30% of threatened Red List species in Australia and surrounding regions.

Power plants, factories, and vehicles are common sources of air pollution. In some cases, the pollutants are directly toxic (for example, lead), but in other cases the pollutants indirectly cause ecological harm when they are present in unnaturally large quantities (for example, carbon dioxide emissions leading to climate change). Not only can air pollutants directly harm animals by causing respiratory issues and cancer as well as damage vegetation, but some interact with the atmosphere to form **acid deposition** (commonly called acid rain). Acid deposition disrupts aquatic ecosystems as well as soil communities and plant growth.

Heavy metals, plastics, pesticides, herbicides, fertilizers, and sediments are examples of water pollution. Heavy metals (including copper, lead, mercury, and zinc) can leach into soil and water from mines. Furthermore, **acid mine drainage** is caused by reaction of mine wastes, such as sulfides, with rainfall or groundwater to produce acids, like sulfuric acid. The Environmental Protection Agency estimates that 40% of the watersheds in the western United States are contaminated by mine run-off. Plastics harm shorebirds, turtles and aquatic invertebrates that ingest and accumulate them. Nutrients, such as nitrate and phosphates, are healthy in bodies of water to an extent, but when fertilizer pollution adds too many of these nutrients at one time, algal blooms can result. This has cascading effects that can ultimately shade and kill aquatic plants and deplete oxygen needed by fish and other animals (**eutrophication**). A particularly concerning water pollution problem is **micropollutants**. For examples, some chemical residues affect growth, cause birth defects, and have other toxic effects on humans and other organisms even at very low concentrations.

Figure $4.2.5.a$ summarizes the effects of air and water pollution on biodiversity, and the chapters about Solid Waste Management, Water Pollution, and Air Pollution explain these threats in more detail.

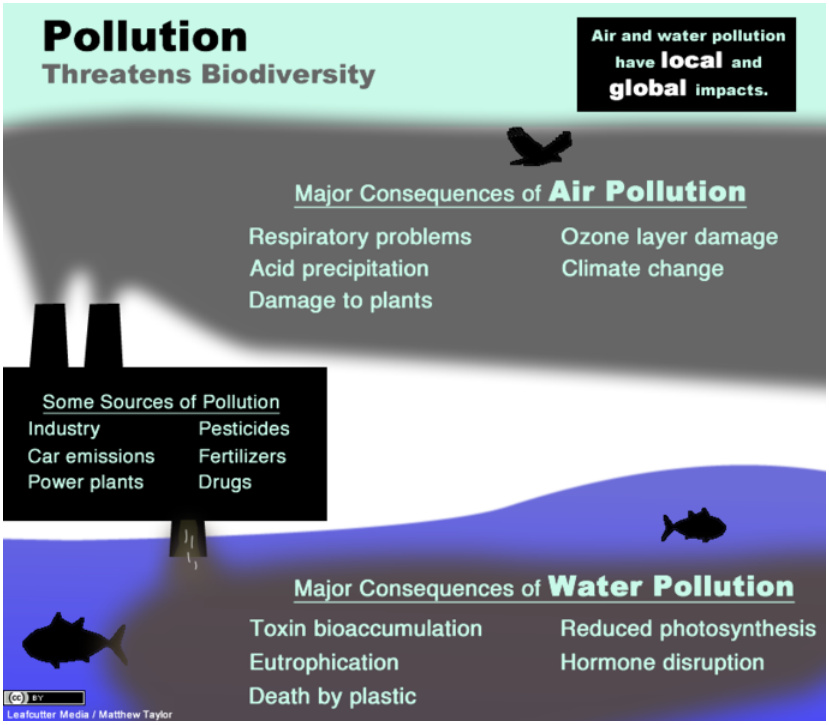


Figure 4.2.5.*a*: Some sources of water and air pollution and their major effects on biodiversity. Sources of pollution include industry, car emissions, power plants, pesticides, fertilizers, and drugs. Some major consequences of air pollution include respiratory problems, acid precipitation, plant damage, ozone depletion, and climate change. Water pollution can bioaccumulate and can lead to eutrophication, death by plastic, reduced photosynthesis, and hormone disruption. Image by Leafcutter Media (CC-BY-SA).

References

Kingsford RT, Watson JEM, Lundquist CJ, Venter O, Hughes L, Johnston EL, Atherton J, Gawel M, Keith DA, Mackey BG, Morley C, Possingham HP, Raynor B, Recher HF, Wilson KA. Major conservation policy issues for biodiversity in Oceania. Conservation Biology 2009;23(4):834–40.

Attributions

Modified by Melissa Ha from the following sources:

- Issues and Opinions from *AP Environmental Science* by University of California College Prep, University of California (licensed under CC-BY). Download for free at CNX.
- Biodiversity 2011- Pollution and Biodiversity 2016- Pollution from Australia State of the Environment by Commonwealth of Australia (licensed under CC-BY)

This page titled 4.2.5: Pollution is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **10.5: Pollution** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

4.2.6: Invasive Species

Invasive species are non-native organisms that, when introduced to an area out of its native range, disrupt the community they invade. **Non-native** (exotic) refers to species occurring outside of their historic distribution. Invasive species have been intentionally or unintentionally introduced by humans into an ecosystem in which they did not evolve. Human transportation of people and goods, including the intentional transport of organisms for trade, has dramatically increased the introduction of species into new ecosystems. These new introductions are sometimes at distances that are well beyond the capacity of the species to ever travel itself and outside the range of the species' natural predators. Invasive species can cause ecological and economic damage. They threaten other species through competition for resources, predation, or disease.

In the United States, invasive species like the purple loosestrife (*Lythrum salicaria*) and the zebra mussel (*Dreissena polymorpha*) have drastically altered the ecosystems they invaded. Some well-known invasive animals include the emerald ash borer (*Agrilus planipennis*) and the European starling (*Sturnus vulgaris*; figure 4.2.6.a). Whether enjoying a forest hike, taking a summer boat trip, or simply walking down an urban street, you have likely encountered an invasive species.

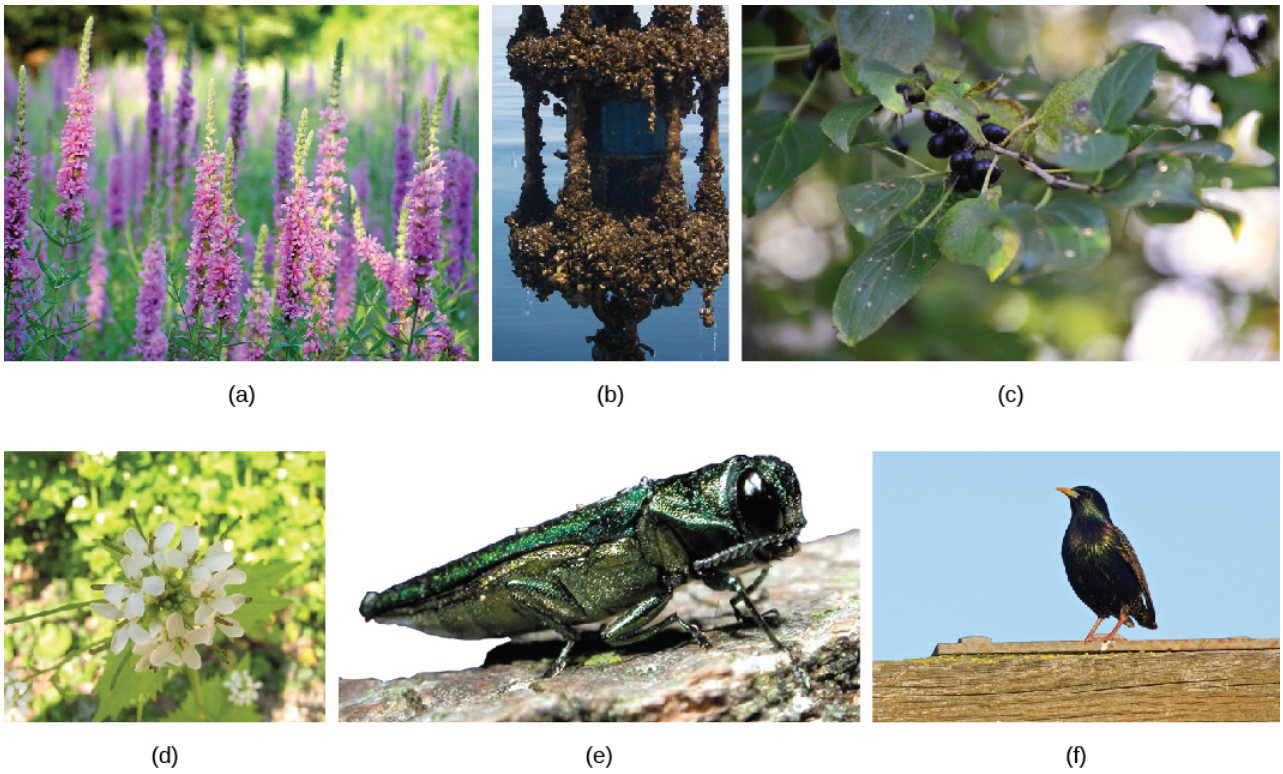


Figure 4.2.6.a: In the United States, invasive species like (a) purple loosestrife (*Lythrum salicaria*) and the (b) zebra mussel (*Dreissena polymorpha*) threaten certain aquatic ecosystems. Some forests are threatened by the spread of (c) common buckthorn (*Rhamnus cathartica*), (d) garlic mustard (*Alliaria petiolata*), and (e) the emerald ash borer (*Agrilus planipennis*). The (f) European starling (*Sturnus vulgaris*) may compete with native bird species for nest holes. (credit a: modification of work by Liz West; credit b: modification of work by M. McCormick, NOAA; credit c: modification of work by E. Dronkert; credit d: modification of work by Dan Davison; credit e: modification of work by USDA; credit f: modification of work by Don DeBold)

Asian Carp

One of the many recent proliferations of an invasive species concerns the Asian carp in the United States. Asian carp were introduced to the United States in the 1970s by fisheries (commercial catfish ponds) and by sewage treatment facilities that used the fish's excellent filter feeding abilities to clean their ponds of excess plankton. Some of the fish escaped, and by the 1980s they had colonized many waterways of the Mississippi River basin, including the Illinois and Missouri Rivers.

Voracious feeders and rapid reproducers, Asian carp may outcompete native species for food and could lead to their extinction. One species, the grass carp, feeds on phytoplankton and aquatic plants. It competes with **native** species (those that historically occurred in the area and are adapted to the local ecosystem) for these resources and alters habitats for other fish by removing aquatic plants.

In some parts of the Illinois River, Asian carp constitute 95 percent of the community's biomass. Although edible, the fish is bony and not desired in the United States.

The Great Lakes and their prized salmon and lake trout fisheries are being threatened by Asian carp. The carp are not yet present in the Great Lakes, and attempts are being made to prevent its access to the lakes through the Chicago Ship and Sanitary Canal, which is the only connection between the Mississippi River and Great Lakes basins. To prevent the Asian carp from leaving the canal, a series of electric barriers have been used to discourage their migration; however, the threat is significant enough that several states and Canada have sued to have the Chicago channel permanently cut off from Lake Michigan. Local and national politicians have weighed in on how to solve the problem. In general, governments have been ineffective in preventing or slowing the introduction of invasive species.

Effect on Endemic Species

Lakes and islands are particularly vulnerable to extinction threats from introduced species. In Lake Victoria, the intentional introduction of the Nile perch was largely responsible for the extinction of about 200 species of cichlids (see Patterns of Biodiversity). The accidental introduction of the brown tree snake via aircraft (figure 4.2.6.*b*) from the Solomon Islands to Guam in 1950 has led to the extinction of three species of birds and three to five species of reptiles endemic to the island. Several other species are still threatened. The brown tree snake is adept at exploiting human transportation as a means to migrate; one was even found on an aircraft arriving in Corpus Christi, Texas. Constant vigilance on the part of airport, military, and commercial aircraft personnel is required to prevent the snake from moving from Guam to other islands in the Pacific, especially Hawaii. Islands do not make up a large area of land on the globe, but they do contain a disproportionate number of endemic species because of their isolation from mainland ancestors.

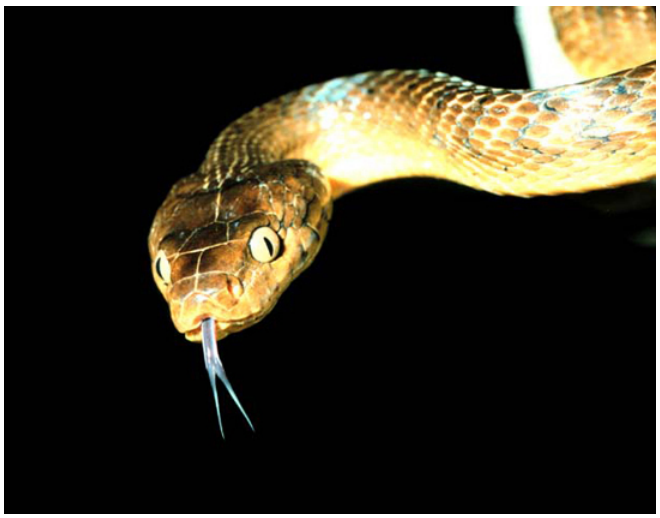


Figure 4.2.6.*b*: The brown tree snake, *Boiga irregularis*, is an exotic species that has caused numerous extinctions on the island of Guam since its accidental introduction in 1950. (credit: NPS)

Introduction by Ballast Water

Many introductions of aquatic species, both marine and freshwater, have occurred when ships have dumped ballast water taken on at a port of origin into waters at a destination port. Water from the port of origin is pumped into tanks on a ship empty of cargo to increase stability. The water is drawn from the ocean or estuary of the port and typically contains living organisms such as plant parts, microorganisms, eggs, larvae, or aquatic animals. The water is then pumped out before the ship takes on cargo at the destination port, which may be on a different continent. The zebra mussel was introduced to the Great Lakes from Europe prior to 1988 in ballast water. The zebra mussels in the Great Lakes have created millions of dollars in clean-up costs to maintain water intakes and other facilities. The mussels have also altered the ecology of the lakes dramatically. They threaten native mollusk populations, but have also benefited some species, such as smallmouth bass. The mussels are filter feeders and have dramatically improved water clarity, which in turn has allowed aquatic plants to grow along shorelines, providing shelter for young fish where it did not exist before. The European green crab, *Carcinus maenas*, was introduced to San Francisco Bay in the late 1990s, likely in

ship ballast water, and has spread north along the coast to Washington. The crabs have been found to dramatically reduce the abundance of native clams and crabs with resulting increases in the prey species of those native crabs.

Invasive Species as Diseases

Invading exotic species can also be disease organisms. It now appears that the global decline in amphibian species recognized in the 1990s is, in some part, caused by the fungus *Batrachochytrium dendrobatidis* (Bd), which causes the disease chytridiomycosis (figure 4.2.6.*c*). There is evidence that the fungus is native to Africa and may have been spread throughout the world by transport of a commonly used laboratory and pet species: the African clawed frog, *Xenopus laevis*. It may well be that biologists themselves are responsible for spreading this disease worldwide. The North American bullfrog, *Rana catesbeiana*, which has also been widely introduced as a food animal but which easily escapes captivity, survives most infections of *B. dendrobatidis* and can act as a disease **reservoir** by storing the infectious fungus.



Figure 4.2.6.*c*: This Limosa harlequin frog (*Atelopus limosus*), an endangered species from Panama, died from a fungal disease called chytridiomycosis. The red lesions are symptomatic of the disease. (credit: Brian Gratwicke)

Early evidence suggests that another fungal pathogen, *Geomyces destructans*, introduced from Europe is responsible for white-nose syndrome, which infects cave-hibernating bats in eastern North America and has spread from a point of origin in western New York State (figure 4.2.6.*d*). The disease has decimated bat populations and threatens extinction of species already listed as endangered: the Indiana bat, *Myotis sodalis*, and potentially the Virginia big-eared bat, *Corynorhinus townsendii virginianus*. How the fungus was introduced is unknown, but one logical presumption would be that recreational cavers unintentionally brought the fungus on clothes or equipment from Europe.



Figure 4.2.6.*d*: This little brown bat in Greeley Mine, Vermont, March 26, 2009, was found to have white-nose syndrome. (credit: modification of work by Marvin Moriarty, USFWS).

Biological Control of Invasive Species

One reason why invasive species proliferate dramatically outside of their native range is due to **release from predators**. This means that parasites, predators, or herbivores that usually regulate their populations are not present, allowing them to outcompete or overpredate native species, which are still regulated. Based on this principle, organisms that regulate the invasive species populations have been introduced to the newly colonized areas in some cases. The release of organisms (or viruses) to limit population size is called **biological control**. As described in the examples below, biological control of invasive species has had varying success, exacerbating the problem in some cases and solving it in others.

Prickly-pear Cactus (*Opuntia*)

Introduced into Australia, this cactus soon spread over millions of hectares of range land driving out forage plants. In 1924, the cactus moth, *Cactoblastis cactorum*, was introduced (from Argentina) into Australia. The caterpillars of the moth are voracious feeders on prickly-pear cactus, and within a few years, the caterpillars had reclaimed the range land without harming a single native species. However, its introduction into the Caribbean in 1957 did not produce such happy results. By 1989, the cactus moth had reached Florida, and now threatens five species of native cacti there.

Purple Loosestrife

The leaf beetle (*Galerucella calmariensis*) has been introduced to suppress purple loosestrife, a noxious weed (figure 4.2.6.*e*). A combination of four biological controls, including the leaf beetle were released in Minnesota since 1992. While it has not eradicated populations of this invasive species, biological control largely removed leaves from 20% of the purple loosestrive populations where it was released, which could reduce competition for native species. The biological controls established populations in most locations where they were released and even spread to new patches of purple loosestrife.



Figure 4.2.6.*e*: Young larvae of the leaf-beetle feed in and on the developing buds of plants, often destroying them. This may stunt plant growth and delay or prevent flowering. Adults (shown) and older larvae feed on leaves and cause severe defoliation. Leaf-beetles can be used to as a biocontrol for invasive plants such as purple loosestrife.

Klamath Weed

In 1946 two species of *Chrysolina* beetles were introduced into California to control the Klamath weed (St. Johnswort) that was ruining millions of acres of range land in California and the Pacific Northwest. Before their release, the beetles were carefully tested to make certain that they would not turn to valuable plants once they had eaten all the Klamath weed they could find. The beetles succeeded beautifully, restoring about 99% of the endangered range land and earning them a commemorative plaque at the Agricultural Center Building in Eureka, California.

European Rabbit

In 1859, the European rabbit was introduced into Australia for sport. With no important predator there, it multiplied explosively (figure 4.2.6.*f*). The raising of sheep (another imported species) suffered badly as the rabbits competed with them for forage.

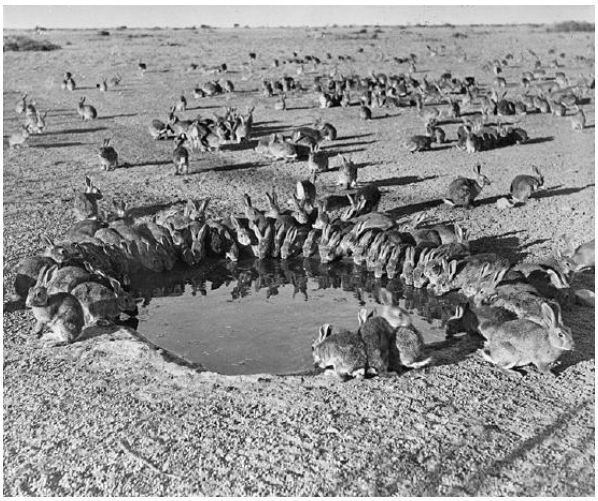


Figure 4.2.6.*f*: These rabbits in Australia removed all forage plants, which ordinarily supply them with water as well as food. They thus had to drink from a pool. Image by National Archives of Australia (public domain).

In 1950, the **myxoma virus** was brought from Brazil and released. The epidemic that followed killed off millions of rabbits (more than 99% of the population). Green grass returned, and sheep raising once again became profitable. Rabbit populations gradually increased, however, because the rabbits evolved to be more resistant to the virus, and the myxoma virus evolved to cause less damage. (Parasites, like viruses, benefit from multiplying inside the host and spreading to other individuals. If they kill their hosts too soon, they typically limit opportunities to multiply and spread.) More recently, the rabbit hemorrhagic disease virus has been used as biological control.

Strategies for Effective Biological Control

To summarize the lessons learned from biological control successes and failures, only candidates that have a very narrow target preference (eat only a sharply-limited range of hosts) should be chosen. Each candidate should be carefully tested to be sure that once it has cleaned up the intended target, it does not turn to desirable species. Biological controls must not be used against native species. Finally, introduction of non-native species into the environment should be avoided because they could themselves be invasive.

Attributions

Modified by Melissa Ha from the following sources:

- Threats to Biodiversity, Community Ecology, and Sustainable Agriculture from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Threats to Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)
- Biological Control and Symbiosis from *Biology* by John W. Kimball (licensed under CC-BY)

This page titled 4.2.6: Invasive Species is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **10.6: Invasive Species** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

4.2.7: Climate Change

Global **climate change** is also a consequence of human population needs for energy, and the use of fossil fuels to meet those needs. Essentially, burning **fossil fuels**, including as oil, natural gas, and coal, increases carbon dioxide concentrations in the atmosphere. Carbon dioxide traps heat energy from the sun, resulting not only in an average increase in global temperature but also in changing precipitation patterns and increased frequency and severity of extreme weather events, such as hurricanes. Scientists overwhelmingly agree the present warming trend is caused by humans. See the Climate Change chapter for a detailed description its cause and impacts. A few examples of how climate change impacts biodiversity are described in the paragraphs below.

Climate change is recognized as a major extinction threat, particularly when combined with other threats such as habitat loss. Scientists disagree about the likely magnitude of the effects, with extinction rate estimates ranging from 15 percent to 40 percent of species committed to extinction by 2050. By altering regional climates, it makes habitats less hospitable to the species living in them. The warming trend will shift colder climates toward the north and south poles, forcing species to move (if possible) with their adapted climate norms. For example, one study indicates that European bird species ranges have moved 91 kilometers (56.5 miles) northward, on average. The same study suggested that the optimal shift based on warming trends was double that distance, suggesting that the populations are not moving quickly enough. Range shifts have also been observed in plants, butterflies, other insects, freshwater fishes, reptiles, amphibians, and mammals.

The shifting ranges will impose new competitive regimes on species as they find themselves in contact with other species not present in their historic range. One such unexpected species contact is between polar bears and grizzly bears (figure $4.2.7.a$). Previously, these two species had separate ranges. Now, their ranges are overlapping and there are documented cases of these two species mating and producing viable offspring.



Figure 4.2.7.*a*: Historically, grizzly bear habitat extended from Mexico through the western United States and into the mid-latitudes of Canada. Since 2008, grizzly bears (*Ursus arctos horribilis*) have been spotted farther north than their historic range, a possible consequence of climate change. Their range now extends to the northern tip of Canada and throughout Alaska. As a result, grizzly bear habitat now overlaps polar bear (*Ursus maritimus*) habitat. The two kinds of bears, which are capable of mating and producing viable offspring, are considered separate species as historically they lived in different habitats and never met. However, in 2006 a hunter shot a wild grizzly-polar bear hybrid known as a grolar bear, the first wild hybrid ever found.

Climate gradients will also move up mountains, eventually crowding species higher in altitude and eliminating the habitat for those species adapted to the highest elevations. Some climates will completely disappear. The rate of warming appears to be accelerated in the arctic, which is recognized as a serious threat to polar bear populations that require sea ice to hunt seals during the winter months. Seals are a critical source of protein for polar bears. A trend to decreasing sea ice coverage has occurred since observations began in the mid-twentieth century. The rate of decline observed in recent years is far greater than previously predicted by climate models.

Changing climates also throw off the delicate timing adaptations that species have to seasonal food resources and breeding times. Scientists have already documented many contemporary mismatches to shifts in resource availability and timing. For example, pollinating insects typically emerge in the spring based on temperature cues. In contrast, many plant species flower based on daylength cues. With warmer temperatures occurring earlier in the year, but daylength remaining the same, pollinators ahead of peak flowering. As a result, there is less food (nectar and pollen) available for the insects and less opportunity for plants to have their pollen dispersed. For migrating birds, timing is everything – they must arrive at their summer breeding grounds when food supplies are at their peak, so that they can rebuild their body fat and reproduce successfully. In some areas, birds are showing up early, before flowers open or insects hatch, and finding very little to eat.

Ocean levels rise in response to climate change due to meltwater from glaciers and the greater volume occupied by warmer water. Shorelines will be inundated, reducing island size, which will have an effect on some species, and a number of islands will disappear entirely. Additionally, the gradual melting and subsequent refreezing of the poles, glaciers, and higher elevation mountains—a cycle that has provided freshwater to environments for centuries—will be altered. This could result in an overabundance of salt water and a shortage of freshwater.

Finally, increased carbon dioxide levels in the atmosphere react with ocean water to form carbonic acid, a phenomenon called **ocean acidification**. In combination with warmer temperatures, ocean acidification is responsible for coral bleaching, the process by which coral expel the algae that typically conduct photosynthesis within the corals. Ocean acidification can also dissolve the calcium carbonate skeletons formed by the coral. Overall, climate change plays a major role in the loss of nearly one third of coral reefs.

Attributions

Modified by Melissa Ha from the following sources:

- Threats to Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Threats to Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)
- First Order Effects from *AP Environmental Science* by University of California College Prep, University of California (licensed under CC-BY). Download for free at CNX.

This page titled 4.2.7: Climate Change is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **10.7: Climate Change** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

4.2.8: Data Dive- Threats to Biodiversity

Overview

Starting in 1998, the World Wildlife Federation (WWF) began publishing the Living Planet Report. This report analyzes the health of the plant relative to the impact humanity has on it, and is updated every two years. This report continually urges humanity, especially world leaders, to work together to build a more sustainable, resilient, and healthy future for people and nature. In the 2018 update, there was an overview of the major threats to biodiversity for several animal families of concern. See the graph of their results below:

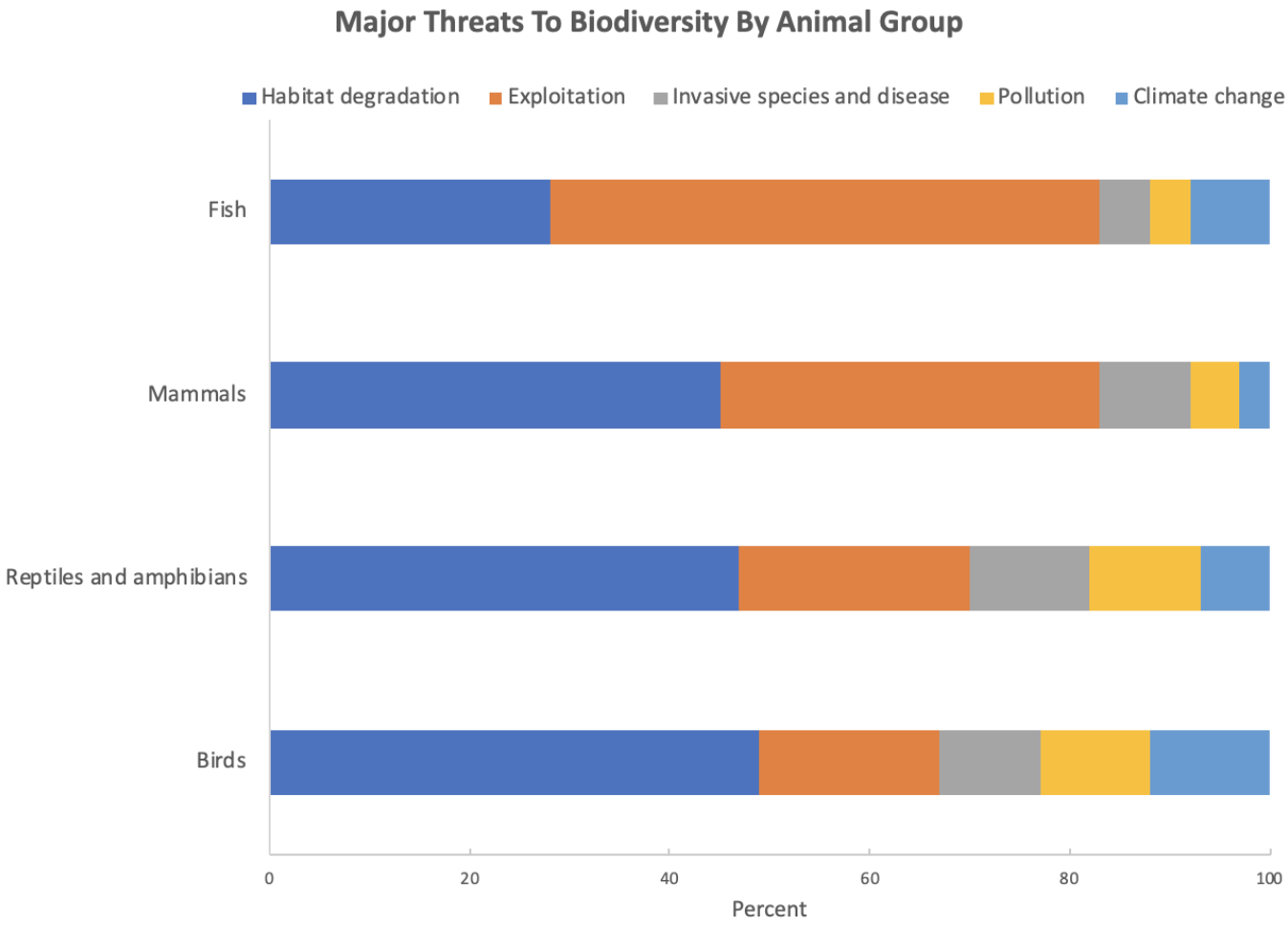


Figure 4.2.8.a: Percent of biodiversity threats to various animal groups. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Living Planet Report 2018 Aiming Higher.

Questions

1. What is the independent (explanatory) variable and the dependent (response) variable?
2. What animal category above is most affected by habitat destruction? What about exploitation?
3. What category above has the largest impact (on average) on all species?
4. What is different for the fish data patterns compared to the other animal categories? Why?
5. How can the results of this graph to inform future conservation efforts?
6. Based on what you have learned in the chapter, list two other categories of life that you think should be added to the graph. Provide at least one reason for each as to why you think it should be added.

Raw Data From Above Graph(s)

Table 4.2.8.*a***:** Raw data for percent of biodiversity threats to various animal groups. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Living Planet Report 2018 Aiming Higher.

Animal Group	Habitat Degradation	Exploitation	Invasive Species And Disease	Pollution	Climate Change
Birds	49	18	10	11	12
Reptiles/Amphibians	47	23	12	11	7
Mammals	45	38	9	5	3
Fish	28	55	5	4	8

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 4.2.8: Data Dive- Threats to Biodiversity is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **10.8: Data Dive- Threats to Biodiversity** by Melissa Ha and Rachel Schleiger has no license indicated.

4.2.9: Review

Summary

After completing this chapter you should be able to...

- Describe biodiversity as the equilibrium of naturally fluctuating rates of extinction and speciation.
- Summarize potential causes of mass extinctions and the associated biodiversity loss.
- Compare the present rate of extinction to the background extinction rate.
- Describe the causes and consequences of the sixth mass extinction.
- Describe how the loss of biodiversity is measured.
- Distinguish among the Red List categories.
- Name, define, and provide examples of the five major threats to biodiversity.
- Provide examples of the successes and failures of biological control in regulating invasive species.

Five mass extinctions with losses of more than 50 percent of extant species are observable in the fossil record. The **sixth mass extinction** is currently in progress with present-day extinction rates much greater than the **background extinction rate**.

The core threats to biodiversity are human population growth and unsustainable resource use. These are habitat loss, overexploitation, pollution, invasive species, and climate change. **Habitat loss** occurs through deforestation, damming of rivers, and other activities. Water and air **pollution** introduce toxic substances into the environment that harm plants and animals. **Overexploitation** is a threat particularly to aquatic species, but the poaching of terrestrial animals and overcollection of plants and fungi also puts species at risk. **Invasive species** have been the cause of a number of extinctions and are especially damaging to islands and lakes. **Climate change** is forcing range changes that can lead to extinction. It is also affecting adaptations to the timing of resource availability that negatively affects species in seasonal environments. Climate change will also raise sea levels, eliminating some islands and reducing the area of all others.

Attributions

Modified by Melissa Ha from the following sources:

- Conservation & Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- The Biodiversity Crisis and Threats to Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)

This page titled 4.2.9: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **10.9: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

4.3: Protecting Biodiveristy

Chapter Hook

As of 2020, there were 41,415 species on the International Union for Conservation of Nature (IUCN) Red List, 16,306 of them being endangered and threatened with extinction. This is up from 16,118 from 2019. The Island Fox (*Urocyon littoralis*) is only found on six of the California Channel Islands off the coast of southern California. In the mid-1990's four island fox subspecies experienced drastic population declines. Soon after, they were listed as critically endangered on the IUCN red list and endangered on the U.S Endangered Species Act. However, due to uncompromising recovery efforts (through supported funding from its endangered status), all four subspecies have almost completely recovered their populations. Regrettably, not all species that end up on the IUCN red list or endangered species list have trajectories that later hold such hope.



Figure 4.3.*a* Island fox (*Urocyon littoralis*). Image by USFWS (licensed under CC-BY 2.0)

The field of **conservation** focuses on preserving biodiversity. Effective conservation depends on ecological knowledge. Today, the main efforts to preserve biodiversity involve legislative approaches to regulate human and corporate behavior, setting aside protected areas, and ecosystem restoration.

Attribution

Modified by Rachel Schleiger and Melissa Ha from Preserving Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

4.3.1: Policies

4.3.2: Non-profit Organizations

4.3.3: Species-level Conservation

4.3.4: Protected Areas

4.3.5: Ecosystem Restoration

4.3.6: Economic Influences on Conservation

4.3.7: Individual Choices

4.3.8: Data Dive- Island Fox Populations

4.3.9: Review

This page titled 4.3: Protecting Biodiveristy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

4.3.1: Policies

National and State Laws

Within many countries there are laws that protect endangered species and regulate hunting and fishing. For example, the **Endangered Species Act** (ESA) was enacted in 1973 in the United States. The ESA does not automatically protect species categorized as threatened on the Red List. Instead, the U.S. Fish and Wildlife Service (FWS), which enforces the ESA, assesses candidates for protected status as threatened or endangered (figure 4.3.1.a). Consideration of candidate species can be initiated by the FWS itself or at the request of the public. Through the Species Status Assessment Framework, the FWS compiles biological data, such as habitat and population information and current threats to the species. This biological data is used to inform decisions.



Figure 4.3.1.a: Public Affairs Specialist for the U.S. Fish and Wildlife Service, Joanna Gilkeson, takes photos of federally endangered San Diego fairy shrimp in a vernal pool in Otay Mesa, California. Vernal pools are shallow seasonal ponds, and vernal pool species are uniquely adapted to changes in water availability. Image by Maideline Sanchez/USFWS (public domain).

Once a species is listed, the FWS is required by law to develop a management plan to protect the species and bring it back to sustainable numbers. The ESA, and others like it in other countries, is a useful tool, but it suffers because it is often difficult to get a species listed or to get an effective management plan in place once a species is listed.

The 1972 **Marine Mammals Protection Act** prohibits the "take" of marine mammals—including harassment, hunting, capturing, collecting, or killing—in U.S. waters and by U.S. citizens on the high seas. The act also makes it illegal to import marine mammals and marine mammal products into the United States without a permit.

State laws can also aid in conservation. Through **California Endangered Species Act** (CESA), originally passed in 1970 and subsequently amended, the California Fish and Game Commission assesses species to be listed as threatened or endangered by the state. A listed species, or any part or product of the plant or animal, may not be imported into the state, exported out of the state, "taken" (killed), possessed, purchased, or sold without proper authorization.

Note that endangered is a subcategory of threatened on the Red List, but for ESA and CESA, threatened and endangered are separate categories, with the latter representing at the greater risk of extinction.

International Agreements

The **Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES)** treaty came into force in 1975. The treaty, and the national legislation that supports it, provides a legal framework for preventing listed species from being transported across nations' borders, thus protecting them from being caught or killed when the purpose involves international trade.

Species can be listed in one of three CITES appendices. Trade is banned for Appendix I species, which are threatened with extinction. For example the trade of ivory is banned by CITES. Trade is regulated for Appendix II species, such as the meat and shells of the queen conch or big-leaf mahogany (figure 4.3.1.*b*). Appendix III species are protected in at least one country, and the local government needs a coordinated response through CITES. For example, several species of red and pink corals have been added to Appendix III at the request of China.

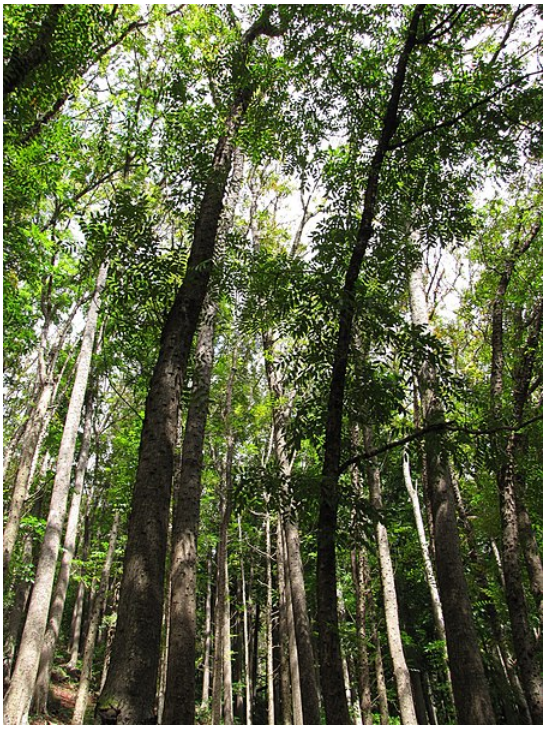


Figure 4.3.1.*b*: Big-leaf mahogany (*Swietenia macrophylla*) timber is used to make furniture, musical instruments, boats, and other products. It listed in CITES Appendix II, meaning that it is regulated. Specifically, a permit is required to ship big-leaf mahogany timber. Image by Forest and Kim Starr (CC-BY).

Approximately 35,800 species are protected by the CITES. The treaty is limited in its reach because it only deals with international movement of organisms or their parts. It is also limited by various countries' ability or willingness to enforce the treaty and supporting legislation.

The **Migratory Bird Treaty Act** (MBTA) is an agreement between the United States and Canada that was signed into law in 1918 in response to declines in North American bird species caused by hunting. The Act now lists over 800 protected species. It makes it illegal to disturb or kill the protected species or distribute their parts (much of the hunting of birds in the past was for their feathers). Examples of protected species include northern cardinals, the Red-tailed Hawk, and the American Black Vulture.

Attributions

Modified by Melissa Ha from the following sources:

- Preserving Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Threatened and Endangered Species by California Department of Fish and Wildlife (public domain)
- Marine Mammal Protection by NOAA Fisheries (public domain)

This page titled 4.3.1: Policies is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **11.1: Policies** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

4.3.2: Non-profit Organizations

The private non-profit sector plays a large role in the conservation effort both in North America and around the world. For example, the World Wildlife Fund raised over $300 million in the 2017-2018 fiscal year to protect endangered animals, their habitats, and natural resources. Some non-profit organizations are species specific, like the Orangutan Foundation, and others are broadly focused, such as the IUCN and Trade Records Analysis of Flora and Fauna in Commerce (TRAFFIC). Illegal wildlife trade is monitored by TRAFFIC. The Nature Conservancy (figure 4.3.2.a) takes a novel approach. It purchases land and protects it in an attempt to set up preserves for ecosystems.



Figure 4.3.2.a: Jo-Ellen Darcy (left), Assistant Secretary of the Army for Civil Works, discusses the Hamilton City with Ryan Luster, project director for The Nature Conservancy, during a site tour in Hamilton City, Calif., June 3, 2014. The U.S. Army Corps of Engineers Sacramento District's Hamilton City project will let nature help do the work of reducing flood risk by restoring 1,500 acres of native habitat between the Sacramento River and a new setback levee to be built near the town. The Nature Conservancy assisted the project partners by helping acquire the land planned for restoration. Image by US Army Corps of Engineers (public domain).

Attributions

Modified by Melissa Ha from the following sources:

- Preserving Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)
- Changing Human Behavior in Response to Biodiversity Loss from *General Biology* by Boundless (CC-BY-SA)

This page titled 4.3.2: Non-profit Organizations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **11.2: Non-profit Organizations** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

4.3.3: Species-level Conservation

Some conservation efforts center around a single species. Often this is a charismatic species that elicits public interest, such as tigers, sea otters, or the California Condor. The specific approach depends on the specific threats to the focus species. When California Condor population size plummeted, leaving only 23 individuals, they were relocated into a controlled environment and provided with optimal conditions for reproduction. This is called **captive breeding**. Individuals have since been **reintroduced** into the environment. One threat to the species was lead from bullets that were entering the food chain and ultimately poisoning the California Condor. Public education and providing lead-free bullets was thus another component of their conservation plan. By 2016, the population size increased to 446 individuals, with 276 of those living in the wild.

When disease puts species at risk, vaccination may be part of the conservation plan. For example, the black-footed ferret of the Great Plains is threatened by sylvatic plague as well as habitat loss. In response, peanut buttered-flavored oral vaccines were distributed over their habitats. Conservation efforts also focused on captive breeding and reintroduction. Once thought to be extinct, there are now several hundred black-footed ferrets in captivity and similar numbers in the wild (figure 4.3.3.a).

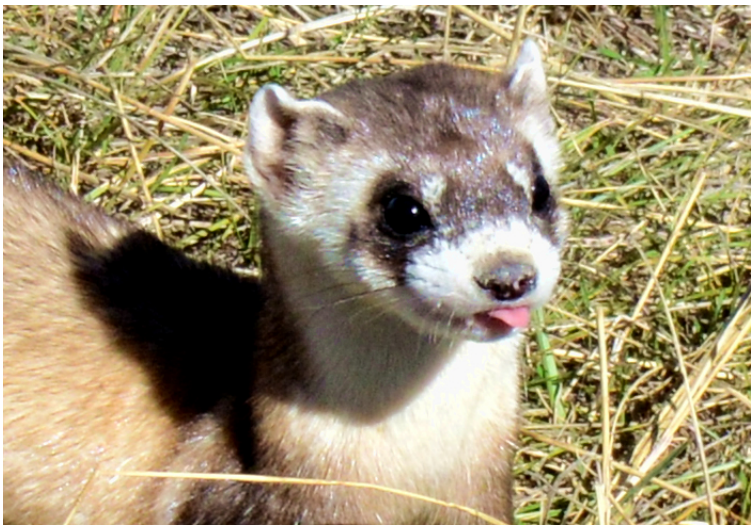


Figure 4.3.3.a: A black-footed ferret kit at the National Black-footed Ferret Conservation Center in Colorado. Image by Ryan Moehring/USFWS (CC-BY)

Protecting or restoring habitat is another component of species-level conservation. For example, the Northern Spotted Owl lives in old-growth forests. Protecting intact old-growth forests and restoring forest habitat is thus critical to their success. This is accompanied by removal of the Barred Owl, an invasive that competes with the Northern Spotted Owl. Similarly, removal of the invasive iceplant *Carpobrotus edulis* from California coasts, restores conditions for endangered dune vegetation (figure 4.3.3.b-d). In other cases, habitat restoration involves promoting populations of native species that support those at risk, such as promoting vegetation that provides food and shelter for waterfowl.



Figure 4.3.3.*b*: Vegetation at Martin Dunes in California. Iceplant is visible in the forefront. Photo by USFWS (public domain).



Figure 4.3.3.*c*: *Carpobrotus edulis* is an invasive iceplant that forms thick mats and outcompetes coastal vegetation. Image by Dr. Dwayne Meadows, NOAA/NMFS/OPR (public domain).



Figure 4.3.3.*d*: Removal of the iceplant restores habitat for endangered species, such as Menzies' wallflower. Image by Gordon Leppig & Andrea J. Pickart, FWS (public domain).

Seemingly unimpressive species can still serve vital ecological roles. For example, the delta smelt is an important food source for larger fish species. Additionally, it serves as an **indicator species** because the health of its populations reflect overall ecosystem health. While the delta smelt has been a focus of species-level conservation, non-charismatic species are often overlooked. In fact, a 2007 study by Colléony and colleagues found that people more often donated to conservation efforts for species that were more similar to humans rather than choosing those that were at greatest risk of extinction. Broad approaches such as establishing protected areas and ecosystem restoration benefit charismatic and non-charismatic species alike. Additionally, broad approaches protect unidentified and species that have not been assessed.

References

Colleony, Agathe, Susan Clayton, Denis Couvet, Michel Saint Jalme, and Anne-Caroline Prevot. (2017). Human preferences for species conservation: Animal charisma trumps endangered status. *Biological Conservation*. 206. 263-269. DOI

Attribution

Melissa Ha (CC-BY-NC)

This page titled 4.3.3: Species-level Conservation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **11.3: Species-level Conservation** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

4.3.4: Protected Areas

It is important to protect natural areas for several reasons. Some people feel a cultural or spiritual connection to the wilderness. Every year, millions of people visit recreational lands such as parks and wilderness areas to experience attractions of the great outdoors: hiking among the giant sequoias in California, traveling on a photo safari in Kenya or just picnicking at a local county park. Besides providing people with obvious health benefits and aesthetic pleasures, recreational lands also generate considerable tourist money for government and local economies. Outdoor recreation activities such as hiking and camping benefit tourist industries and manufacturers of outdoor clothes and equipment.

Establishment of preserves is one of the key tools in conservation efforts. A **preserve** is an area of land set aside with varying degrees of protection for the organisms that exist within the boundaries of the preserve (figure 4.3.4.a). Governments or private organizations establish nature preserves. In 2016, the IUCN estimated that 14.7 percent of Earth's land surface was covered by preserves of various kinds. This area is large, but only 20% of the key biodiversity areas identified by the IUCN were sufficiently protected.



Figure 4.3.4.a: Mequon Nature Preserve in Wisconsin. Image by Jennifer Tomaloff (CC-BY-NC-SA).

There has been extensive research into optimal preserve designs for maintaining biodiversity. Preserves can be seen as "islands" of habitat within "an ocean" of non-habitat. In general, large preserves are better because they support more species, including species with large home ranges; they have more core area of optimal habitat for individual species; they have more niches to support more species; and they attract more species because they can be found and reached more easily. One large preserve is better than the same area of several smaller preserves because there is more core habitat unaffected by less hospitable ecosystems outside the preserve boundary. For this same reason, preserves in the shape of a square or circle will be better than a preserve with many thin "arms." If preserves must be smaller, then providing wildlife corridors between two preserves is important so that species and their genes can move between them. All of these factors are taken into consideration when planning the nature of a preserve before the land is set aside. In addition to the physical specifications of a preserve, there are a variety of regulations related to the use of a preserve. These can include anything from timber extraction, mineral extraction, regulated hunting, human habitation, and nondestructive human recreation.

The public lands described below differ in their level of protection. For example, national parks and forests allow camping whereas wildlife refuges place more limitations on human activities.

Wilderness Areas

Wilderness areas, comprise ecosystems in which human activity has not significantly affected the plant and animal populations or their environment. Natural processes predominate. According to the "Wilderness Act of 1964," wilderness areas are defined as being those areas where the nearest road is at least five miles away and where no permanent buildings stand. Activities that could

disrupt native species, such as the use of motorized vehicles is prohibited. More than 100 million acres of land are now preserved as wilderness under this act. Sparsely populated Alaska contains the largest chunk of wilderness areas, over half of it. Although wilderness areas are scattered among most of the lower 48 states, the largest percentage is found in the western states. Few undesignated areas in the contiguous states remain that would qualify as wilderness.

National parks and forests and wildlife refuges can contain wilderness areas. California contains significant wilderness areas, with over 4 million acres of National Forest Wilderness areas, and 1.5 million acres of mostly desert wilderness in the Mojave Desert National Preserve (figure 4.3.4.*b*).



Figure 4.3.4.*b*: A variety of unique plant life grows at Mojave Desert National Preserve. Image by John Fowler (CC-BY).

Wilderness areas provide an essential habitat for a wide array of fish, wildlife, and plants, and are particularly important in protecting endangered species. For scientists, wilderness areas serve as natural laboratories, where studies can be performed that would not be possible in developed areas.

National and State Parks

The United States has set aside more land for public recreational use than any other country. The **National Park System** manages more than 380 parks, recreation areas, seashores, trails, monuments, memorials, battlefields, and other historic sites. It consists of more than 80 million acres nationwide (figure 4.3.4.*c*). The largest national park is Wrangell–St. Elias National Park and Preserve in Alaska with over 13 million acres. California has eight national parks: Channel Islands, Death Valley, Joshua Tree, Lassen, Redwood, Sequoia, Kings Canyon, and Yosemite. Many national parks such as Yosemite, Yellowstone and the Grand Canyon are such popular recreation destinations that the ecosystems of those parks are being severely tested by human activities.



Figure 4.3.4.*c*: National parks, such as Grand Teton National Park in Wyoming, help conserve biodiversity. (credit: Don DeBold)

Every state has also set aside significant amounts of land for recreational use. The **California State Park System** manages more than one million acres of parklands including: coastal wetlands, estuaries, scenic coastlines, lakes, mountains and desert areas. California's largest state park is Anza-Borrego Desert State Park, which is the largest state park in the United States with 600,000 acres. The stated mission of the California State Park System is "to provide for the health, inspiration and education of the people

of California by helping to preserve the state's extraordinary biological diversity, protecting its most valued natural and cultural resources and creating opportunities for high-quality outdoor recreation".

National Forests

The **National Forest System**, managed by the U.S. Forest Service (part of the United States Department of Agriculture), consists of more than 170 forestlands and grasslands, which are available for activities such as camping, fishing, hiking and hunting. These are managed as multiple use lands, which balance the needs for recreation, grazing, timber, watershed protection, wildlife and fish, and wilderness.

Some examples of national forests are the Sierra National Forest in California and the White Mountain National Forest in New Hampshire. The Coronado National Forest in Arizona is famous for "sky islands", or steep mountain ranges surrounded by low-lying areas. The dramatic increase in elevation is associated with changes in the flora and fauna (figure 4.3.4.*d*). Explore national forests using this interactive map.



Figure 4.3.4.*d*: The Santa Teresa Mountains of the Coronado National Forest form "sky islands". Image by Jstuby (public domain).

Wildlife Refuges

The U.S. Fish and Wildlife Service manage more than 500 national **wildlife refuges** (figure 4.3.4.*e*), which not only protect animal habitats and breeding areas but also provide recreational facilities. Find a Refuge is an interactive map for locating wildlife refuges.



Figure 4.3.4.*e*: Bobcat at Sonny Bono National Wildlife Refuge in southern California. Image by Mark Stewart/USFWS (public domain).

Supplemental Reading

America's Public Lands Explained. 2016. U.S. Department of the Interior.

Attributions

Modified by Melissa Ha from the following sources:

- Preserving Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Preserving Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)
- Biological and Land from *AP Environmental Science* by University of California College Prep, University of California (licensed under CC-BY). Download for free at CNX.
- Multiple Use Lands. 12 May 2015. The United States Department of Justice. Accessed 28 March 2021 (public domain)

This page titled 4.3.4: Protected Areas is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **11.4: Protected Areas** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

4.3.5: Ecosystem Restoration

Ecosystem restoration is the process of bringing an area back to its natural state, before it was impacted through destructive human activities. It holds considerable promise as a mechanism for maintaining or restoring biodiversity and reinstating ecosystem services. It requires a broad interdisciplinary approach involving many different scientific fields of study (for example, biology, ecology, hydrology and geology). Reintroducing wolves, a top predator, to Yellowstone National Park in 1995 led to dramatic changes in the ecosystem that increased biodiversity. The wolves (figure 4.3.5.*a*) function to suppress elk and coyote populations and provide more abundant resources to the detritivores. Reducing elk populations has allowed revegetation of **riparian areas** (those along the banks of a stream or river), which has increased the diversity of species in that ecosystem. Reduction of coyote populations by wolves has increased the prey species previously suppressed by coyotes. In this ecosystem, the wolf is a **keystone species**, meaning a species that is instrumental in maintaining diversity within an ecosystem. Removing a keystone species from an ecological community causes a collapse in diversity. The results from the Yellowstone experiment suggest that restoring a keystone species effectively can have the effect of restoring biodiversity in the community. Ecologists have argued for the identification of keystone species where possible and for focusing protection efforts on these species. It makes sense to return the keystone species to the ecosystems where they have been removed.

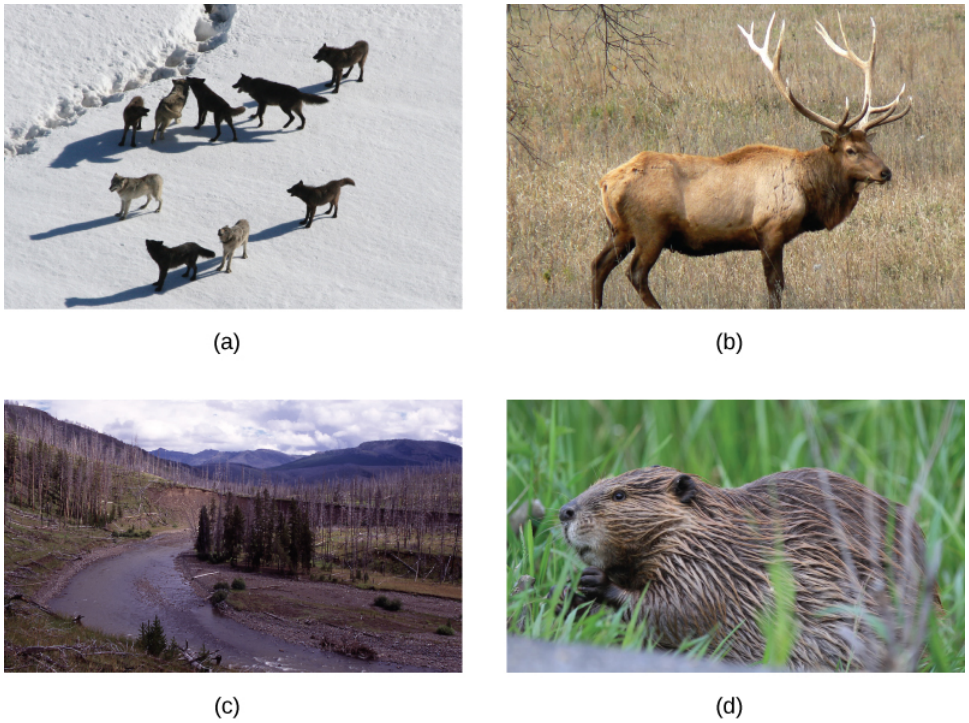


Figure 4.3.5.*a*: (a) The Gibbon wolf pack in Yellowstone National Park, March 1, 2007, represents a keystone species. The reintroduction of wolves into Yellowstone National Park in 1995 led to a change in the grazing behavior of (b) elk. To avoid predation, the elk no longer grazed exposed stream and riverbeds, such as (c) the Lamar Riverbed in Yellowstone. This allowed willow and cottonwood seedlings to grow. The seedlings decreased erosion and provided shading to the creek, which improved fish habitat. A new colony of (d) beaver may also have benefited from the habitat change. (credit a: modification of work by Doug Smith, NPS; credit c: modification of work by Jim Peaco, NPS; credit d: modification of work by "Shiny Things"/Flickr)

Other large-scale restoration experiments underway involve dam removal. In the United States, since the mid-1980s, many aging dams are being considered for removal rather than replacement because of shifting beliefs about the ecological value of free-flowing rivers. The measured benefits of dam removal include restoration of naturally fluctuating water levels (often the purpose of dams is to reduce variation in river flows), which leads to increased fish diversity and improved water quality (see Dams and Reservoirs for more about the impact of dams). In the Pacific Northwest of the United States, dam removal projects are expected to increase populations of salmon, which is considered a keystone species because it transports nutrients to inland ecosystems during its annual spawning migrations. In other regions, such as the Atlantic coast, dam removal has allowed the return of other spawning anadromous fish species (species that are born in fresh water, live most of their lives in salt water, and return to fresh water to spawn). Some of the largest dam removal projects have occurred recently, such as Elwha Dam on the Olympic Peninsula of

Washington State. The large-scale ecological experiments that these removal projects constitute will provide valuable data for other dam projects slated either for removal or construction.

Besides physical processes, socioeconomic factors must also be considered in a restoration project. Actions of humans have historically been important in shaping ecosystems, and are important in determining the success of restoration efforts. Because the cost to restore an individual site can involve millions of dollars, government support is a necessity.

Environmental Remediation

Danger to human health from both historic and modern pollution requires that cleanup measures be implemented. **Remediation** is aimed at neutralization, containment, and/or removal of the polluting chemicals. The goal is to prevent the spread of the pollution, or to reduce it to levels that will not appreciably risk human health. Many times, it is physically impossible or financially unfeasible to completely clear all contamination. Often, experts and the public disagree on how clean is clean enough.

Many communities are struggling to find the funds and technological expertise needed to clean up polluted areas. Some settings, such as brownfields, can be remediated fairly easily. **Brownfields** are abandoned industrial or commercial facilities or blighted urban areas that need to be cleansed of contamination before they can be redeveloped. Other areas, because of their size or the extreme toxicity of their contaminants, require very expensive, complex, and long-term remediation. Many of these have been designated as Superfund sites.

Superfund sites are areas with the most toxic contamination in the United States. The contamination may not only make the site itself too dangerous to inhabit, but often leaks toxic levels of pollutants into the surrounding soil, water, or air. An example of a Superfund site is Love Canal in Niagara Falls, New York (figure 4.3.5.*b*). The canal was a chemical waste dump for many years, then in the 1950's was covered with soil and sold to the city. Over time, many homes and a school were built over the former dump. In the 1970's, heavy rains raised the water table and carried contaminants back to the surface. Residents noticed foul smells, and gardens and trees turned black and died. Soon after, rates of birth defects, cancer, and other illnesses began to rise sharply. In 1977, the State of New York and the federal government began remediation work. Buildings were removed, and all residents were bought out and relocated, contaminated deposits and soils were excavated, and remaining soils and groundwater were treated and sealed off to prevent further spread of the contamination. Remediation activities have now been completed at this site.



Figure 4.3.5.*b*: Love Canal. Source: US Environmental Protection Agency

The type of pollution and the medium affected (air, water, or soil) determine remediation methods. Methods include incineration, absorption onto carbon, chemical methods, or bioremediation. **Bioremediation** is the use of plants, bacteria, or fungus to "digest" the contaminant to a non-toxic or less toxic form. All of these methods tend to be expensive and time-consuming.

Reclamation and Mitigation

Reclamation involves salvaging some features of a degraded habitat, but it may not restore the ecosystem fully (figure 4.3.5.*c*). For example, instead of abandoning a mined area once resources have been collected, it can be reclaimed by planting vegetation, reshaping the landscape, and redirecting water flow. However the reclaimed land still lacks many features of the original ecosystem, such as complex topography, vegetation that took tens or hundreds of years to grow, soil quality, and an intricate network of streams.



Figure 4.3.5.*c*: The Seneca Yoast coal once land has been cleared in preparation for mining (left) and after reclamation (right). Image by Peabody Energy, Inc. (CC-BY).

Sometimes, actions can be taken to avoid, reduce or compensate for the effects of environmental damage. Such **mitigation** efforts have been taken by the Army Corps of Engineers during construction projects. The native plants are removed from a site before construction begins and transplanted at a special holding site. After the construction project is completed, the native plants are replanted using those from the holding site. Another example of mitigation might involve the creation or enhancement of wetlands in one area, in order to compensate for permitted wetland losses in another area. Mitigation often goes hand-in-hand with restoration. Texaco, in conjunction with environmental groups and the United States Fish and Wildlife Service, restored 500 acres of agricultural lands in the lower Mississippi Delta to bottomland hardwoods. Texaco received environmental credits for the mitigating effects of the new woodlands on air quality.

Attributions

Modified by Melissa Ha from the following sources:

- Preserving Biodiversity and Case Study - The Love Canal Disaster from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Preserving Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)
- Issues and Opinions from *AP Environmental Science* by University of California College Prep, University of California (licensed under CC-BY). Download for free at CNX.

This page titled 4.3.5: Ecosystem Restoration is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **11.5: Ecosystem Restoration** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

4.3.6: Economic Influences on Conservation

Economics greatly impacts conservation success. Short-term profits can incentivize individuals, companies, or governments to harvesting resources at an unsustainable rate and at the expense of ecosystem health. In impoverished regions, compromising habitat to grow high-value crops, such as coffee or oil palms, or poaching endangered species may seem like the only source of income. One solution is **debt-for-nature swaps** through which one country forgives the debt of another if the latter agrees to protect natural areas. These conservation efforts can provide a new source of income for residents near the protected areas. For example, the United States forgave $20 million in debt from Costa Rica. In exchange, Costa Rica invested in expanding its protected areas and developing the ecotourism industry, which provides jobs to many of its residents (figure 4.3.6.a). **Ecotourism** involves visiting and enjoying natural areas while minimizing ecological damage. Ecotourism can benefit local economies and alleviate poverty especially if the earnings from it are reinvested into the communities living near tourist destinations. It generates jobs such as park operators, sellers of local crafts, and tour guides.



Figure 4.3.6.a: A tourist ziplines over the Costa Rican rainforest. Image by Khaufle at the English language Wikipedia (CC-BY-SA).

Attribution

Modified by Melissa Ha from Ecotourism from *Life Sciences Grade 10* by Siyavula (licensed under CC-BY)

This page titled 4.3.6: Economic Influences on Conservation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **11.6: Economic Influences on Conservation** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

4.3.7: Individual Choices

Many of the strategies for preserving biodiversity operate at the level of whole governments or large organizations; however, your choices as an individual also play a role in conservation.

Consumer Choices

The products you purchase have differing impacts on biodiversity. Educating yourself on the origin of products and food that you purchase and choosing the sustainable options can help preserve biodiveristy. For example, arabica coffee can be grown in the shade, meaning there is no need to fully clear rainforest vegetation when growing this species (figure 4.3.7.a). However, robusta coffee requires full sun, and cultivating it has a greater impact or rainforest biodiversity. Monterey Bay Aquarium provides a sustainable seafood guide, which identifies seafood choices that have a lower environmental impact. Choosing local products reduces the amount of fossil fuels that were burned to transport them to you, thus reducing carbon emissions that contribute to climate change.



Figure 4.3.7.a: Shade-grown coffee at a Columbian farm. Image by Brian Smith/American Bird Conservancy/U.S. Fish and Wildlife Service Northeast Region (CC-BY).

Some products have special certifications that indicate their impact on biodiversity. For example, certified organic products must be cultivated without the use of synthetic pesticides, herbicides, and fertilizers, which pollute the surrounding areas. Additionally, they cannot contain genetically modified organisms (GMOs), which have both positive and negative environmental implications. If you cannot afford to buy all organic produce, see the Environmental Working Group's lists of the Clean Fifteen™ (for which pesticide use is already limited) and Dirty Dozen™ (which are high priority to purchase organic or avoid due to high pesticide residues). The Roundtable on Sustainable Palm Oil (RSPO) certifies oil palm plantations that follow standards such as avoiding deforestation and using fire to clear land. Additionally, RSPO-certified business must follow guidelines to compensate their employees sufficiently. Fair Trade Certified™ goods must meet social and environmental standards that support the United Nation's Sustainable Development Goals.

Resource Conservation

Resource conservation is individual choice that can promote biodiversity. In this case, conservation refers to limiting one's use of resources, such as water, electricity, and gasoline. Landscaping with drought-tolerant plants to limit the need for irrigation or using a low-flow shower head are examples of water conservation. Because much electricity is generated from burning fossil fuels, such as coal and natural gas, conserving electricity reduces carbon emissions associated with climate change. Turning off lights and appliances when not in use and insulating one's home to reduce electricity spent on heating and cooling save money and benefit the environment. Similarly, transportation choices such as carpooling, biking, or taking public transportation can limit carbon emissions. Reusing items or not purchasing unnecessary ones also conserves the energy needed to produce and transport these

goods and reduces plastic waste, which is particularly harmful to aquatic ecosystems. See chapters about Water Resources, Renewable Energy, and Solid Waste Management for more about resource conservation.

Guide to Planting a Pollinator Garden

Whether you have a few feet on your apartment balcony or several acres, you can promote populations of native bees, butterflies, and other pollinators by building a **pollinator garden** (figure 4.3.7.*b*). The first step is to choose a location. While flowering plants can grow in both shady and sunny locations, consider your audience. Butterflies and other pollinators like to bask in the sun and some of their favorite wildflowers grow best in full or partial sun with some protection from the wind. The next step is to identify your soil type. Take a look at your soil - is it sandy and well-drained or more clay-like and wet? You can turn over a test patch or check out a soil mapper to learn more. Your soil type and the amount of sunlight it gets will help determine the kinds of plants you can grow.



Figure 4.3.7.*b*: A diversity of plant species support native pollinators in this pollinator garden. Image by Sara "Asher" Morris (CC-BY-NC).

Next, research which varieties of milkweed and wildflowers are native to your area and do well in your soil and sunlight conditions. **Native plants**, those that have historically occurred in the area, are the ideal choice, because they require less maintenance and tend to be heartier. Find a nursery that specializes in native plants near you - they'll be familiar with plants that are meant to thrive in your region. Some examples of pollinator-friendly plants native to California, include the California poppy, California lilac, milkweed, and foothill penstemon. The California Native Plant Society and UC Berkeley Urban Bee Lab have additional plant suggestions. It's essential to choose plants that have not been treated with pesticides. Choosing **perennials** will ensure your plants come back each year, reducing the need for maintenance.

Remember to think about more than just the summer growing season. Pollinators need nectar early in the spring, throughout the summer and even into the fall. Choosing plants that bloom at different times will help you create a bright and colorful garden that both you and pollinators will love for months!

Some native bee species use bare soil for nesting. While applying mulch can help control weeds, leaves some bare soil if possible. Some native pollinators also nest in tiny cavities, which may already occur naturally in or near your garden or can be provided with bee boxes.

Make sure to weed and water your garden to keep it healthy. It may take some time, but you will eventually see butterflies and other pollinators enjoying your garden.

Modified by Melissa Ha from the U.S. Fish and Wildlife Service (public domain).

Citizen Science

Finally, citizen science provides the opportunity to be directly involved in biological conservation efforts. For some opportunities like Globe at Night, which assess light pollution, or the Lost Ladybug Project, data can be collected independently

and submitted online. Others, like bird banding, are scheduled events in which experts train a group of volunteers to complete fieldwork (figure 4.3.7.*c*). The federal government's citizen science database lists many such opportunities.



Figure 4.3.7.*c*: Left: Students returning from taking phenology measurements on native trees in a native forest conservation area on Hawaii Island. Phenology refers to the timing of events in an organism's life cycle. Right: Americorp members preparing ʻōhiʻa tree cuttings for rooting and propagation – with plants to be used for disease resistance screening. Left image by UFS/USDA (public domain). Right image by USFS (public domain).

Attribution

Melissa Ha (CC-BY-NC)

This page titled 4.3.7: Individual Choices is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **11.7: Individual Choices** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

4.3.8: Data Dive- Island Fox Populations

Overview

The Island Fox Conservation Working Group and Friends of the Island Fox nonprofit organization work hard to continually fund research and conservation projects to ensure the continued survival of the island fox. In addition, they also have educational outreach programs about the island fox as an endangered species to get public support. Every June the Island Fox Conservation Working Group meets to go over updates and plan for the future. In the 2019 meeting they discussed population trends across the six island fox subspecies and determine what those populations mean for the subspecies trajectory moving into the future. The results of this meeting can be seen in the graph and table below:

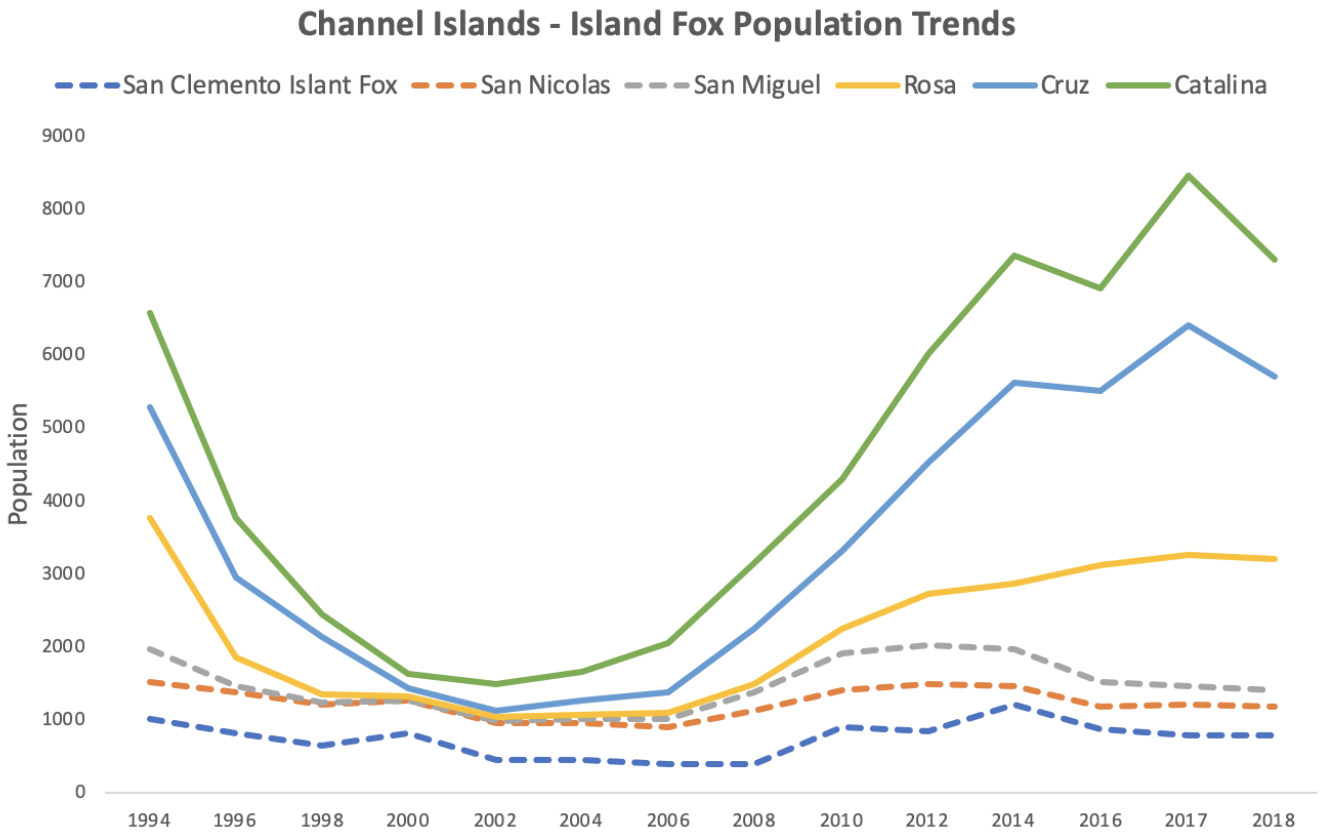


Figure 4.3.8.*a*: Population trends for subspecies of island fox (solid lines = large island subspecies, dashed lines = small island subspecies). Graph by Rachel Schleiger (CC-BY-NC) modified from data in 2019 Island Fox Status Update.

Table 4.3.8.*a***:** The 2019 designated statuses for the subspecies of island fox based on fox population and island size. Graph by Rachel Schleiger (CC-BY-NC) modified from data in 2019 Island Fox Status Update.

Island	Island Size	Fox Population	Status
San Miguel	Small	171	Concerned
San Nicolas	Small	400	Improved
Santa Rosa	Large	1862	Stable
Santa Cruz	Large	2462	Stable
Santa Catalina	Large	1571	Stable
San Clemento	Small	778	Stable

Questions

1. What is the independent (explanatory) variable and the dependent (response) variable?
2. What question(s) are the authors trying to answer with this graph and table?
3. Which subspecies of island fox do you think is the most stable? Why?
4. Which year(s) were bad (meaning population estimates were very low) for most of the island fox subspecies?
5. As noted in the graph, populations are estimated for each subspecies annually. Why do you think an annual estimation was chosen in the conservation plan instead of a longer time interval?

Raw Data From Above Graph(s)

Table 4.3.8.*b***:** Raw data for population trends for subspecies of island fox (solid lines = large island subspecies, dashed lines = small island subspecies). Graph by Rachel Schleiger (CC-BY-NC) modified from data in 2019 Island Fox Status Update.

Year	San Clemento	San Nicolas	San Miguel	Santa Rosa	Santa Cruz	Santa Catalina
1994	1000	520	450	1800	1500	1300
1996	810	550	100	400	1100	800
1998	650	550	40	100	800	300
2000	800	450	20	50	100	200
2002	450	500	25	60	90	350
2004	450	500	50	70	200	375
2006	400	506	100	80	300	650
2008	400	725	260	100	750	900
2010	900	500	500	350	1050	1000
2012	850	640	525	700	1800	1500
2014	1200	263	510	900	2750	1750
2016	860	329	329	1600	2400	1400
2017	775	255	255	1800	3150	2050
2018	790	400	200	1800	2500	1600

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 4.3.8: Data Dive- Island Fox Populations is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **11.8: Data Dive- Island Fox Populations** by Melissa Ha and Rachel Schleiger has no license indicated.

4.3.9: Review

Summary

After completing this chapter you should be able to...

- Describe the legislative framework for conservation, providing and describing examples of national and state laws and international agreements.
- Summarize the role of non-profit organizations in conservation.
- Discuss species-level conservation in terms of its success and shortcomings.
- Provide specific examples of species-level conservation, including the name of the species and the strategies used.
- Explain the importance of protected areas and distinguish between different types of protected areas.
- Describe principles of preserve design.
- Identify examples of the effects of ecosystem restoration, remediation, reclamation, and mitigation.
- Explain how economic factors influence conservation efforts.
- Identify actions that individuals can take to preserve biodiversity.

There is a legislative framework for biodiversity protection. In the United States, the **Endangered Species Act** protects listed species but is hampered by procedural difficulties and a focus on individual species. International treaties such as **CITES** regulate the transportation of endangered species across international borders. The non-profit sector is also very active in funding and organizing conservation efforts.

Conservation involves a variety of approaches, and many factors influence the success of conservation efforts. **Species-level conservation** may include **captive breeding and reintroduction**, vaccinations, and habitat restoration. Presently, 14.7 percent of Earth's land surface is protected in some way. **Preserves** are a major tool in conservation efforts, and large, interconnected preserves favor biodiversity. The U.S. government maintains wilderness areas, national and state parks, national forests, and wildlife refuges, which differ in their levels of protection. **Ecosystem restoration** promotes biodiversity, improves conditions for native species, and reinstates ecosystem services. **Debt-for-nature swaps** and **ecotourism** assist conservation efforts economically. Through choosing products that minimize environmental harm and resource conservation, individuals play an important role in preserving biodiversity.

Attributions

Modified by Melissa Ha from the following sources:

- Conservation & Biodiversity from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Preserving Biodiversity from *General Biology* by OpenStax (licensed under CC-BY)

This page titled 4.3.9: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **11.9: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

CHAPTER OVERVIEW

5: Humans and the Environment

The objective for this unit is to set the stage for understanding how humans have altered the trajectory of Earths abiotic and biotic systems. Specifically this unit will dive into the history of humans on this Earth, how we became successful, how we utilize/acquire basic necessities (water, food), and the how increasing population densities have influenced environmental health. This unit will also addresses the past and present of how the above variables changed as our population changed.

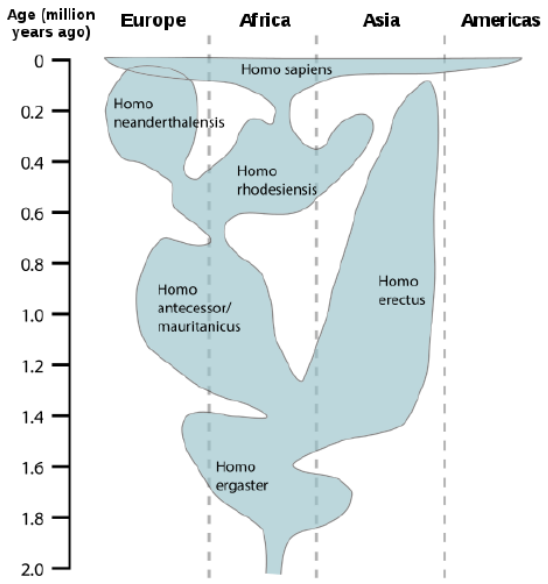


Figure 5.1**:** Chart of human evolution compared to other extinct species of the *Homo* genus. Image by Reed DL, Smith VS, Hammond SL, Rogers AR (licensed under CC-BY 2.5)

Attribution

Rachel Schleiger (CC-BY-NC)

5.1: The Human Population

5.1.1: History of Human Population Growth

5.1.2: The Rate of Human Population Growth

5.1.3: Looking Ahead

5.1.4: Data Dive- World Population Densities

5.1.5: Review

5.2: Water Resources

5.2.1: Fresh Water Supply and the Water Cycle

5.2.2: Water Usage

5.2.3: Water Scarcity and Solutions

5.2.4: Data Dive- Aqueducts of Rome

5.2.5: Review

5.3: Agriculture

5.3.1: Food Security

5.3.2: Industrial Agriculture
5.3.3: Selective Breeding and Genetic Engineering
5.3.4: Sustainable Agriculture
5.3.5: Data Dive- Landslides and Ag Terraces
5.3.6: Review
5.4: Environmental Health
5.4.1: Types of Environmental Hazards
5.4.2: Epidemiology
5.4.3: Infectious Diseases
5.4.4: Environmental Toxicology
5.4.5: Environmental Hazard Reduction
5.4.6: Data Dive- Cholera Cases Worldwide
5.4.7: Review

This page titled 5: Humans and the Environment is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

SECTION OVERVIEW

5.1: The Human Population

Chapter Hook

Why is the human species (*Homo sapiens*) so successful? We have the unique ability to not just occupy, but thrive, in diverse environments, even some of the most extreme. Through our social cooperation (despite large groups with no kin), language, art, and ingenuity (especially with tools and fire), our species became unstoppable. Modern humans have permanent settlements on all but one continent (Antarctica) and have exponentially increased our population. We have changed the world! However, this transformation has consequences for all species (including our own), necessitating updated survival strategies across the board.

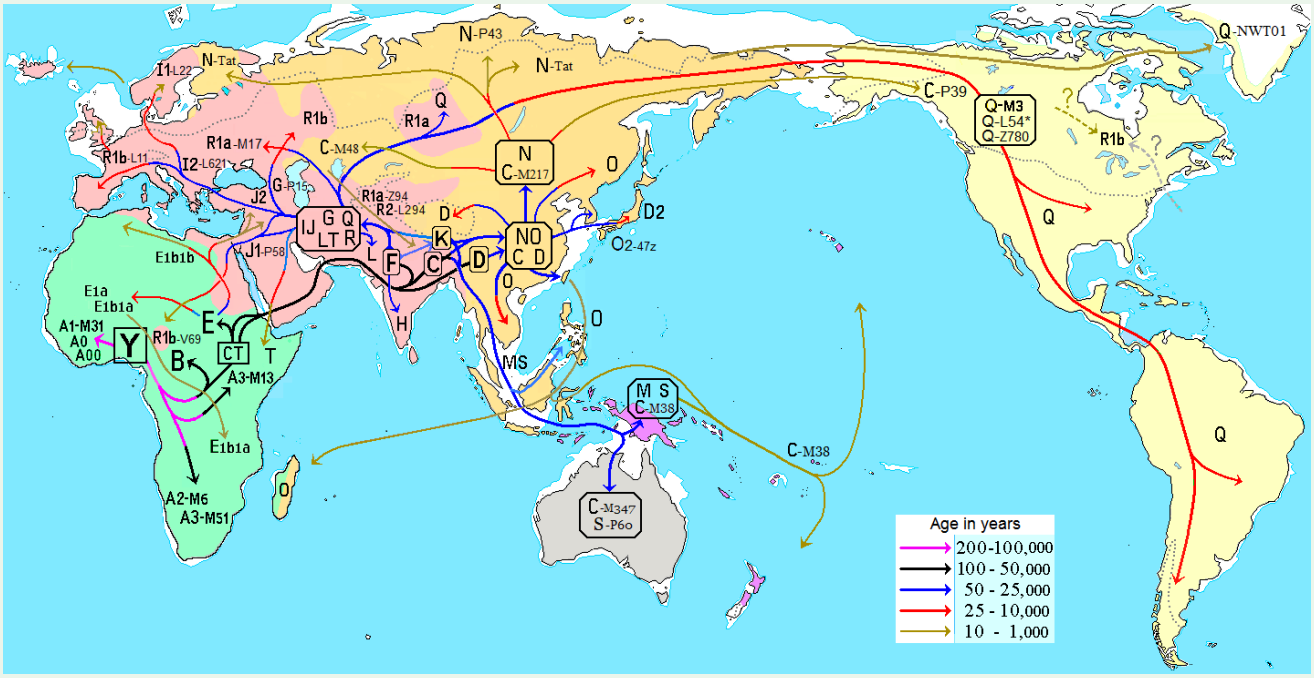


Figure 5.1.*a* Map demonstrating human spread across the globe. Image by Maulucioni (licensed under CC-BY 3.0)

Concepts of population ecology can be applied to human population growth. **Demography** is the study of how human populations grow, shrink, and change in terms of age and gender compositions. Demographers also compare populations in different countries or regions. The human population has increased dramatically in the last few centuries and continues to grow (figure 5.1.*b*). Overpopulation risks human well-being and compromises ecosystem functioning.



Figure 5.1.*b*: The human population continues to grow. China is currently the most populous country with over 1.4 billion people in January of 2022. Image by Sino-German Urbanisation Partnership (public domain).

Attributions

Modified by Rachel Schleiger and Melissa Ha from the following sources:

- Human Demography from *Introduction to Environmental Science*, 2nd edition by Caralyn Zehnder et al. (licensed under CC-BY-NC-SA).
- The Human Population from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

5.1.1: History of Human Population Growth

5.1.2: The Rate of Human Population Growth

5.1.3: Looking Ahead

5.1.4: Data Dive- World Population Densities

5.1.5: Review

This page titled 5.1: The Human Population is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

5.1.1: History of Human Population Growth

The human population is growing rapidly. For most of human history, there were fewer than 1 billion people on the planet. During the time of the Agricultural Revolution, 10,000 B.C., there were only 5-10 million people on Earth - which is basically the population of New York City today. In 1800, during the Industrial Revolution, there were approximately 1 billion people on Earth. Continued agricultural expansion and extraction of fossil fuels and minerals led to rapid global economic growth and, in turn, population growth in the 19th century. We've added over 6 billion people to the human population in just a little over 200 years (figure 5.1.1.*a*). As of August 2020, the global human population is around 7.8 billion people.

Although global population size continues to increase, the rate of human population growth has decreased. This means that the population size is not increasing as quickly as it did in the past (figure 5.1.1.*a*).

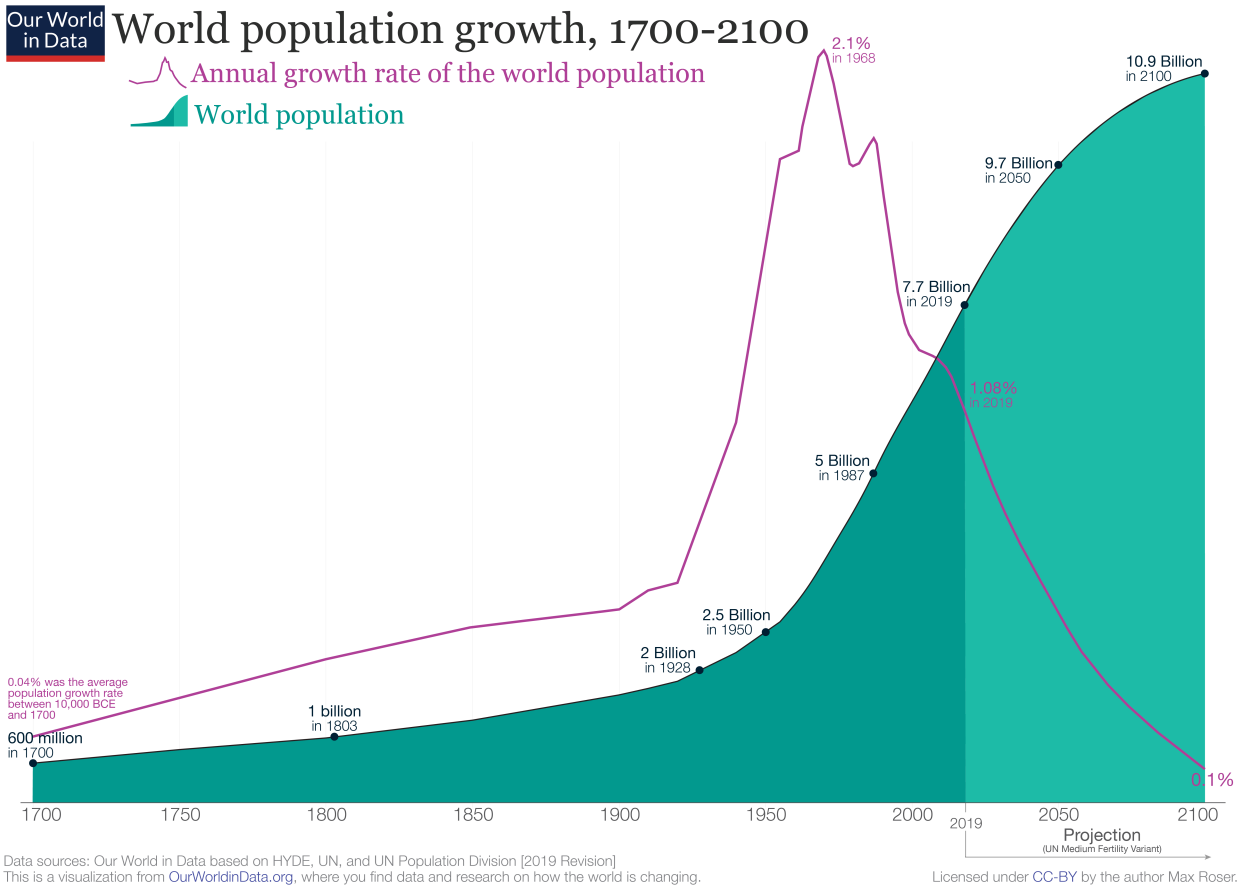


Figure 5.1.1.*a*: Human population growth. Since 1700, human population size (shaded in teal) has increased. It reached 1 billion in 1803, 2 billion in 1928, 2.5 billion in 1950, 5 billion in 1987, and 7.7 billion in 2019. It is projected to reach 10.9 billion in 2100. The population growth rate (pink line) was only 0.04% on average between 10,000 BCE and 1700. The population growth rate peaked in 1968 at 2.1%, and since then, it has slowed to 1.08% in 2019 and is projected to be at 0.1% in 2100. Image by Max Roser (2013) "Future Population Growth". Published online at OurWorldInData.org. (CC-BY)

The fundamental cause of the acceleration of growth rate for humans in the past 200 years has been the reduced death rate due to changes in public health and sanitation. Clean drinking water and proper disposal sewage has drastically improved health in developed nations. Also, medical innovations such as the use of antibiotics and vaccines have decreased the ability of infectious disease to limit human population growth. In the past, diseases such as the bubonic plaque of the fourteenth century killed between 30 and 60 percent of Europe's population and reduced the overall world population by as many as one hundred million people. Naturally, infectious disease continues to have an impact on human population growth, especially in poorer nations. For example, life expectancy in sub-Saharan Africa, which was increasing from 1950 to 1990, began to decline after 1985 largely as a result of HIV/AIDS mortality. According to a 2016 study by Marcus et al., The reduction in life expectancy caused by HIV/AIDS was estimated to be 8 years for 2016.

Human technology and particularly our harnessing of the energy contained in fossil fuels have caused unprecedented changes to Earth's environment, altering ecosystems to the point where some may be in danger of collapse. Changes on a global scale including depletion of the ozone layer, desertification and topsoil loss, and global climate change are caused by human activities.

Reference

Marcus, J. L., Chao, C. R., Leyden, W. A., Xu, L., Quesenberry, C. P., Jr, Klein, D. B., Towner, W. J., Horberg, M. A., & Silverberg, M. J. (2016). Narrowing the Gap in Life Expectancy Between HIV-Infected and HIV-Uninfected Individuals With Access to Care. *Journal of acquired immune deficiency syndromes (1999)*, *73*(1), 39–46.

Attributions

Modified by Melissa Ha from the following sources:

- The Human Population from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Human Demography from *Introduction to Environmental Science*, 2nd edition by Caralyn Zehnder et al. (licensed under CC-BY-NC-SA).
- Human Population Growth from *General Biology* by OpenStax (licensed under CC-BY)
- The Industrialization of Nature: A Modern History (1500 to the present) from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors. Download for free at CNX. (licensed under CC-BY)

This page titled 5.1.1: History of Human Population Growth is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **12.1: History of Human Population Growth** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

5.1.2: The Rate of Human Population Growth

Demographic Transition

Recall from the Populations chapter that the **population growth rate** (*r*) equals the birth rate minus the death rate. Slowly declining birth rates following an earlier sharp decline in death rates are today characteristic of most of the less-developed regions of the world. The shift from high birth and death rates to low birth as well as death rates is called the **demographic transition**.

Prior to World War II, advances in public health were largely limited to affluent, industrialized countries. But since then, many more countries have enjoyed improvements in public health - always with dramatic effect on death rates. For example, in 1945, the death rate in Sri Lanka (then called Ceylon) was 0.022 (2.2%). In 1946, a large-scale program to control mosquitos, which transmit malaria, was started. By eliminating the mosquito, the incidence of malaria dropped sharply. After 9 years, the death rate dropped to 0.010 (1%), and by 2012 was 0.006 (0.6%). However, a compensating decline in birth rates has come more slowly; the birth rate was 0.018 (1.8% in 2012). With birth rates higher than death rates, the population was increasing at an annual rate of 0.012 (1.2%) per year, with a doubling time of 57.5 years (t = 0.69/0.012).

Predicting Future Population Size

Doubling time is based on a constant population growth rate, but this rate can change depending on a number of factors, including total fertility rate (TFR) and the age structure of the population.

Total Fertility Rate

The **total fertility rate** (TFR) is the average number of children that each woman will have during her lifetime. The TFR is an average because, of course, some women will have more, some fewer, and some no children at all. Theoretically, when the TFR = 2, each pair of parents just replaces itself. Actually it takes a TFR of 2.1 or 2.2 to replace each generation - this number is called the **replacement fertility rate** - because some children will die before they grow up to have their own two children. In countries with low life expectancies, the replacement rate is even higher (2.2–3). Figure 5.1.2.*a* compares the total fertility rate in different countries.

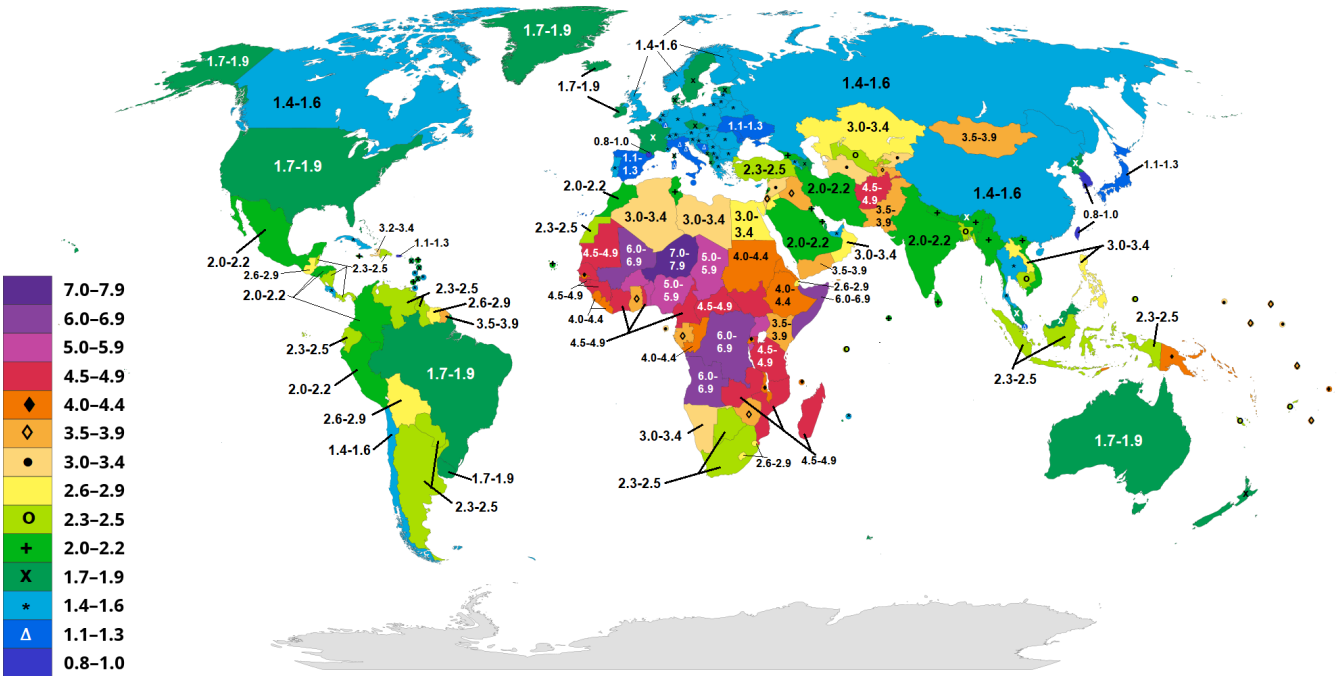


Figure 5.1.2.*a*: Total fertility rate by country in 2020. Most countries in North and South America have a total fertility rate of 1.7-2.5 children per woman, but fertility rates are lower (for example, Canada and Chile) or higher (for example, Guatamala and Bolivia) in a few countries. Most of Europe and northern and Eastern Asia have low fertility rates (generally 1.4-1.6 children per woman). Many Middle Eastern countries and those in southern and southeastern Asia have slightly higher fertility rates (generally 1.7-2.5 children per woman). The fertility rate in Australia is 1.7-1.9 children per woman. High fertility rates are common in African countries. For example, the fertility rate in Niger is 7.0-7.9. Image modified from Korakys (CC-BY-SA).

Age Structure

The **age structure** of a population, the proportion of a population in different age classes, is an important factor in population dynamics. The relationship between TFR and the population growth rate (*r*) depends on age structure. For example, if at one period a population has an unusually large number of children, they will — as they pass through their childbearing years — increase the *r* of the population even if their TFR goes no higher than 2. Most people have children between the ages of 15 and 49. So if a population has a large number of young people just entering their reproductive years, the rate of growth of that population is sure to rise.

Models that incorporate age structure allow better prediction of population growth, plus the ability to associate this growth with the level of economic development in a region. Countries with rapid growth associated with high birth rates have a pyramidal shape in their age structure diagrams, showing a preponderance of younger individuals, many of whom are of reproductive age (figure 5.1.2.*b*).

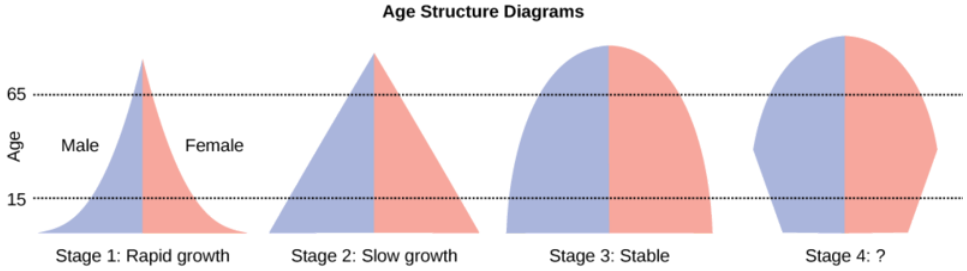


Figure 5.1.2.*b*: Typical age structure diagrams are shown. The rapid growth diagram narrows to a point, indicating that the number of individuals decreases rapidly with age. In the slow growth model, the number of individuals decreases steadily with age. Stable population diagrams are rounded on the top, showing that the number of individuals per age group decreases gradually, and then increases for the older part of the population.

The age structure of a population also reflects the recent pattern of mortality. In countries where injuries, starvation, and disease, etc. take a heavy toll throughout life, the age structure diagram has a broad base. In countries where almost everyone survives until

old age, the bases is narrower.

Rapid growth is most often observed in underdeveloped countries where individuals do not live to old age because of less-than-optimal living conditions, and there is a high birth rate. Age structures of areas with slow growth, including developed countries such as the United States, still have a pyramidal structure, but with many fewer young and reproductive-aged individuals and a greater proportion of older individuals compared to underdeveloped countries. Other developed countries, such as Italy, have zero population growth. The age structure of these populations is more conical, with an even greater percentage of middle-aged and older individuals.The actual growth rates in different countries are shown in figure 5.1.2.*c*, with the highest rates tending to be in the less economically developed countries of Africa and Asia.

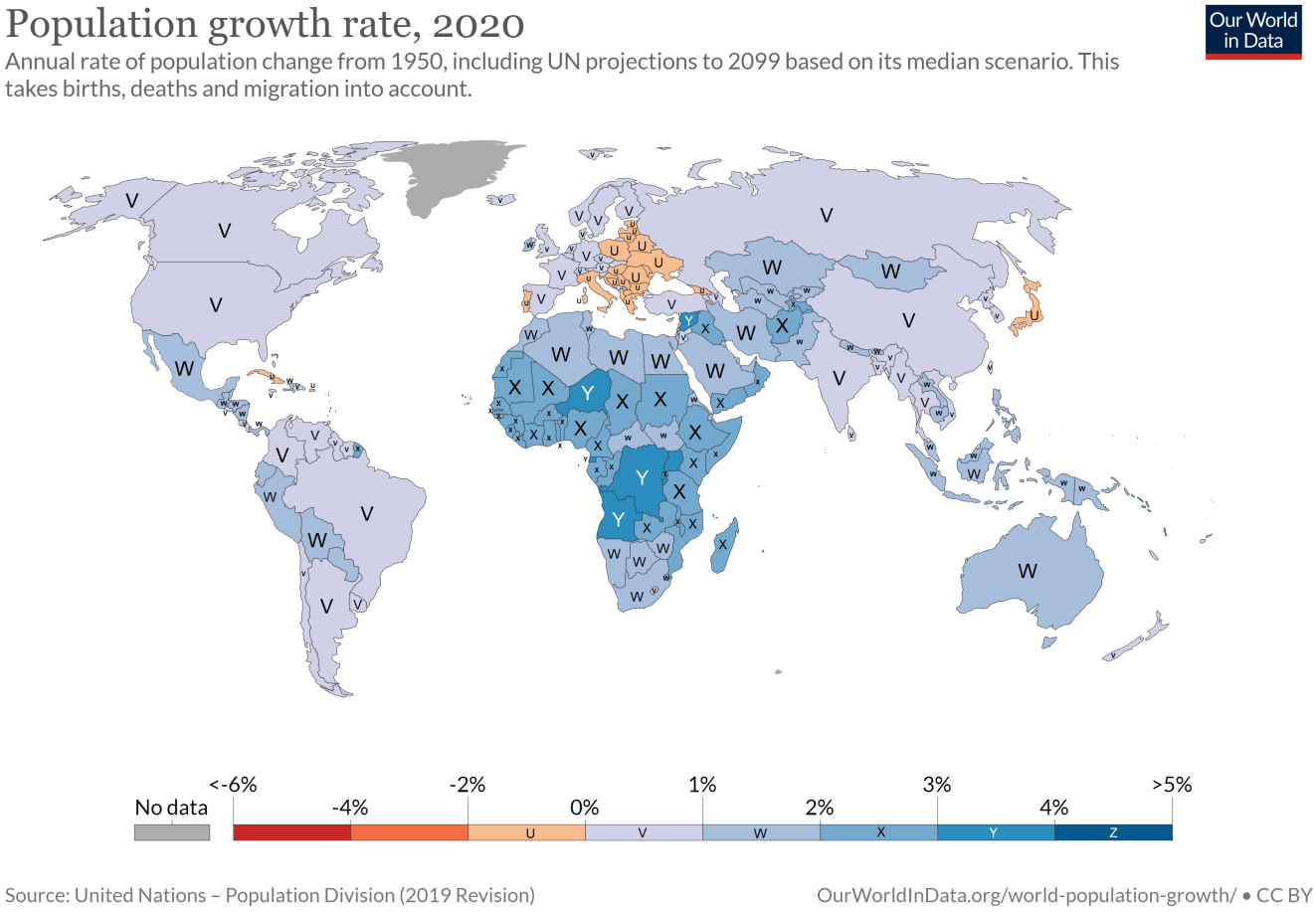


Figure 5.1.2.*c*: The percent growth rate of populations in 2020 in different countries is shown. The population growth rate is negative (meaning the population size is declining) in Cuba, Portugal, many countries in Eastern Europe, and Japan. Mexico, much of Central America, Ecuador, Peru, Bolivia, Paraguay, and Haiti have a population growth rate of 1-2%. Most other countries in North and South America have a population growth rate of 0-1%. Western and Northern Europe and most Asian countries have a population growth rate of 0-1%. In the Middle East, Southeast Asia, Australia, northern and southern Africa, Central African Republic, and South Sudan, the population growth rate is generally 1-2%. Examples of countries with a population growth rate above 2% include the remaining African countries, Syria, Yemen, Oman, Iraq, Afghanistan, and Tajikistan. Image modified from Max Roser, Hannah Ritchie and Esteban Ortiz-Ospina (2013) - "World Population Growth". Published online at OurWorldInData.org (CC-BY).

The U.S. Baby Boom

The TFR in the United States declined from more than 4 late in the nineteenth century to less than replacement in the early 1930s. However, when the small numbers of children born in the depression years reached adulthood, they went on a childbearing spree that produced the baby-boom generation (figure 5.1.2.*d*). In 1957, more children were born in the United States than ever before (or since).

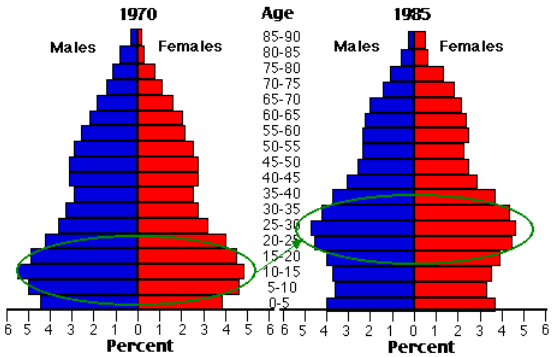


Figure 5.1.2.*d*: Age structure diagrams of the United States population in 1970 and 1985. The x-axis represents the percent of the population. The y-axis represents age groups in increments of five from 0-5 years all the way to 85-90 years. Blue bars to the left represent males, and red bars to the right represent females. Bars taper near the top of the diagram, showing that those in the oldest age categories represent a small percentage of the population. The 1970 age structure diagram bulges around the 10-15 age group, and this represents the baby boom (green oval). In 1985, this bulge has moved to the 25-30 age group.

Attribution

Modified by Melissa Ha from the following sources:

- The Human Population from *Biology* by John W. Kimball (licensed under CC-BY)
- The Human Population from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 5.1.2: The Rate of Human Population Growth is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **12.2: The Rate of Human Population Growth** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

5.1.3: Looking Ahead

Population Projections

A consequence of exponential human population growth is that is takes less time to add a particular number of humans to our Earth. Figure 5.1.3.*a* shows that 123 years were necessary to add 1 billion humans in 1930, but it only took 24 years to add two billion people between 1975 and 1999. As already discussed, at some point it would appear that our ability to increase our carrying capacity indefinitely on a finite world is uncertain. Without new technological advances, the human growth rate has been predicted to continue to slow in the coming decades. However, the population will still be increasing and the threat of overpopulation remains.

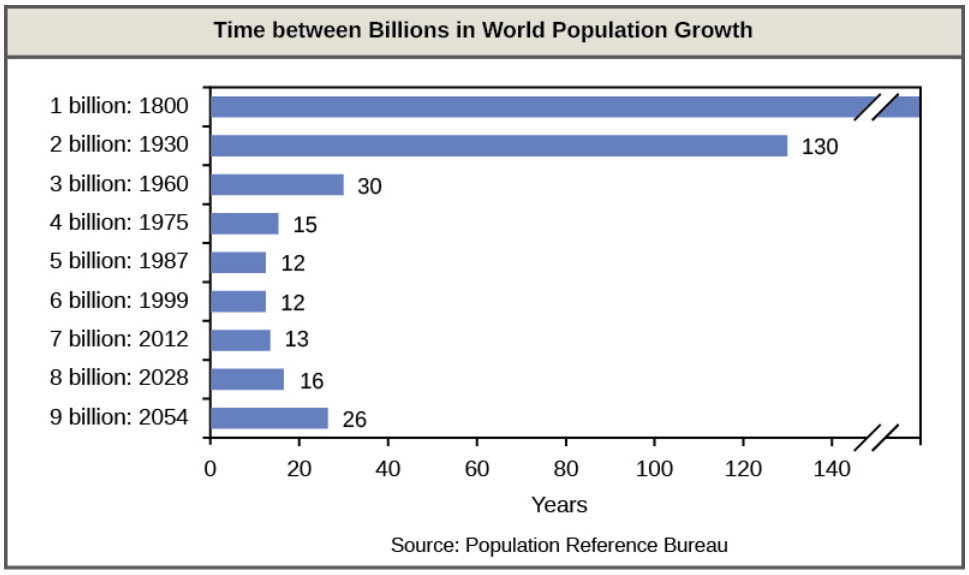


Figure 5.1.3.*a*: The time between the addition of each billion human beings to Earth decreases over time. By 1800, there were about a billion people on Earth. It took 130 years, until 1930, for the number to reach two million. Thirty years later, in 1960, the number reached three billion, and 15 years after that, in 1975, the number reached four billion. The population reached five billion in 1987, and six billion in 1999, each twelve years apart. Currently, the world population exceeds seven billion. The population is projected to reach 8 billion in 2028, and 9 billion in 2054. (credit: modification of work by Ryan T. Cragun)

Exponential growth cannot continue indefinitely. If the current world value for *r* (1.075% in 2019) remains unchanged, the global population would grow from its current 7.8 billion (August 2020) to 9.7 billion by 2050. Experts predict that the global population will continue to grow until 2100, reaching a peak of almost 11 billion. So this begs the question(s)... What is the carrying capacity for humans on Earth? How long until we reach it, or have we already surpassed it? Then finally, what are the consequences of our current and projected populations?

Long-term Consequences

Humans are not unique in their ability to alter their environment. For example, beaver dams alter the stream environment where they are built. Humans, however, have the ability to alter their environment to increase its **carrying capacity**, sometimes to the detriment of other species. Earth's human population and their use of resources are growing rapidly, to the extent that some demographers say that humans have already exceeded our carrying capacity. This would mean that the Earth cannot indefinitely support our current population size. Long-term exponential growth carries with it the potential risks of famine, disease, and large-scale death, as well as social consequences of crowding such as increased crime.

The impact of humans on our environment depends not only on population size but also on resource use. The **ecological footprint** is a measure of the area of land needed to support an individual or country. This includes agricultural lands used to provide food, other lands used to provide goods (such as forests to provide timber) and fuels (such as oil fields to provide fossil fuels), and the land needed to dispose of wastes.

Countries differ in the average ecological footprint per person, and these differences are associated with region (figure 5.1.3.*b*). For example, in 2017, the ecological footprint per person in the United States was 8.1 global hectares compared to only 1.2 global

hectares in India. (Visit footprintnetwork.org to see ecological footprints and related metrics for all countries.) Many factors affect the size of an ecological footprint from economic development to resource conservation to individual product choice. Reducing our ecological footprints can increase our carrying capacity and decrease environmental impact. So maybe we should update the question of, "What the Earths carrying capacity for humans," to, "What is the carrying capacity of people on Earth at which we can provide a fair and decent life for all humans while causing the minimum impact on the environment?

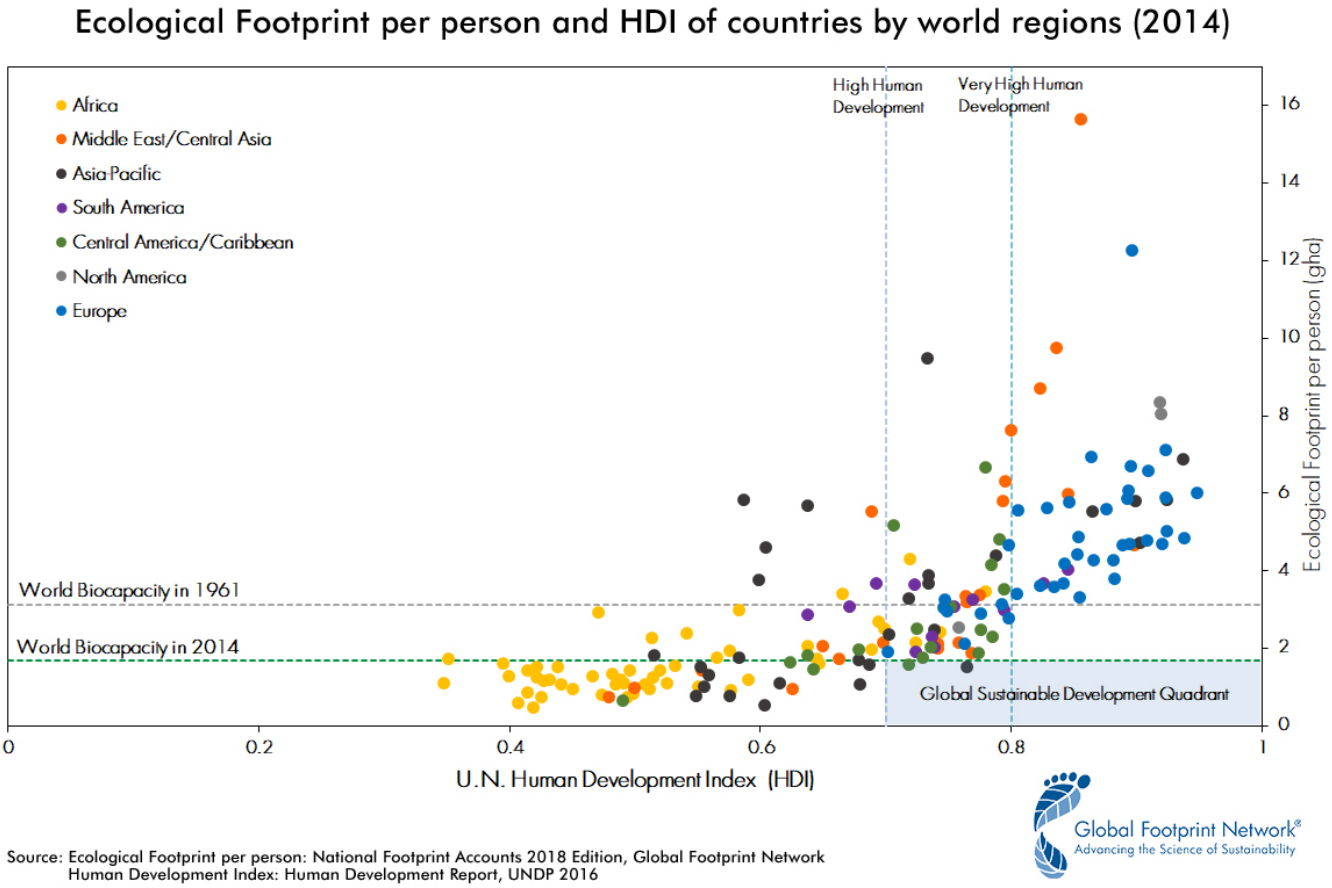


Figure 5.1.3.*b*: The ecological footprint per person in relation to the U.N. human development index for each country coded by region. Countries in some regions of the world have low ecological footprints per person, but this is because people are living in poverty (points in lower left of graph). This pattern is characteristic of many African countries. Other countries have large ecological footprints and live affluently (points in upper right of graph), including many European countries. To sustainably meet human needs, countries must achieve a high human development index and low ecological footprint (points in lower right of graph). Image by Footprint123 (CC-BY-SA).

Are you wondering about the size of your ecological footprint? You can estimate it here.

Regulation

Reducing the population growth rate will limit the consequences of overpopulation; however, regulation of the human population is a complex issue that presents ethical considerations.

Efforts to moderate population control led to the 1979 One-Child Policy in China, which imposed fines on urban couples who had more than one child (figure 5.1.3.*c*). Additionally, only a couple's first child was granted educational and health care benefits. The effectiveness of the policy in limiting overall population growth is controversial, as was the policy itself. Tragically, this policy combined with the desire of many couples to have a male heir led to female infanticide. Furthermore, enforcement of the policy led to human rights abuses including forced sterilizations and forced abortions. Over the years, some restrictions associated with the One-Child Policy were lifted. For example, in 2013, an adult who was an only child was allowed to have two children. The policy ended in 2016, and as of May 2021, each couple is now permitted to have up to three children.



Figure 5.1.3.c: A poster explaining and promoting the One-Child Policy. Image by Arian Zwegers (CC-BY).

Zero population growth occurs when the birth rate equals the death rate such that the rate of population growth is zero. Logically, this would occur when total fertility rate equals replacement fertility rate, but as described in the previous section, this depends on the age structure of a population.

All else being equal, reductions in total fertility rate will slow human population growth. Total fertility rate is negatively correlated with improved standard of living, access to health care, and gender equality. As countries develop economically, individuals become increasingly dependent on urban jobs rather than the family farm for income and provisions. While a large family would be needed to tend the family farm, supporting many children in an urban area is expensive. As a result, economic development is associated with couples choosing to have fewer children.

Access to healthcare decreases childhood mortality and provides the option of birth control. While decreased infant and childhood mortality alone would increase the population growth rate, people tend to have fewer children when they can be confident that all of their children are likely to survive; thus, as childhood mortality declines, total fertility rate also declines, lowering the population growth rate.

The education of women and girls is also associated with a decreased total fertility rate. Furthermore, family planning education programs have had highly positive effects in some countries on limiting population growth rates and increasing standards of living. When women have knowledge of family planning, power to make decisions about their family size, and a variety of career options, fertility rate declines. In regions that have made recent strides in educating girls, families more often choose to have a small number of children so that they can afford to pay for their education regardless of gender.

References

Population Growth (Annual %). 2019 World Bank Group. Accessed 29 August 2020.

Growing at a slower pace, world population is expected to reach 9.7 billion in 2050 and could peak at nearly 11 billion around 2100. 2019. United Nations Department of Economic and Social Affairs. Accessed 29 August 2020.

Attributions

Modified by Melissa Ha from the following sources:

- The Human Population from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- The Human Population from *Biology* by John W. Kimball (licensed under CC-BY)
- Human Population Growth from *General Biology* by OpenStax (licensed under CC-BY)

This page titled 5.1.3: Looking Ahead is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **12.3: Looking Ahead** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

5.1.4: Data Dive- World Population Densities

Overview

The Socioeconomic Data and Applications Center (SEDAC), is one of the Distributed Active Archive Centers (DAACs) in the Earth Observing System Data and Information System (EOSDIS) of the U.S. National Aeronautics and Space Administration (NASA). Its main objective is to investigate human interactions in the environment and to develop and operate applications that support the integration of socioeconomic and earth science data and to serve as an "Information Gateway" between earth sciences and social sciences. One of the projects it continually updates is creating global population density maps. Their 2020 map can be observed below:

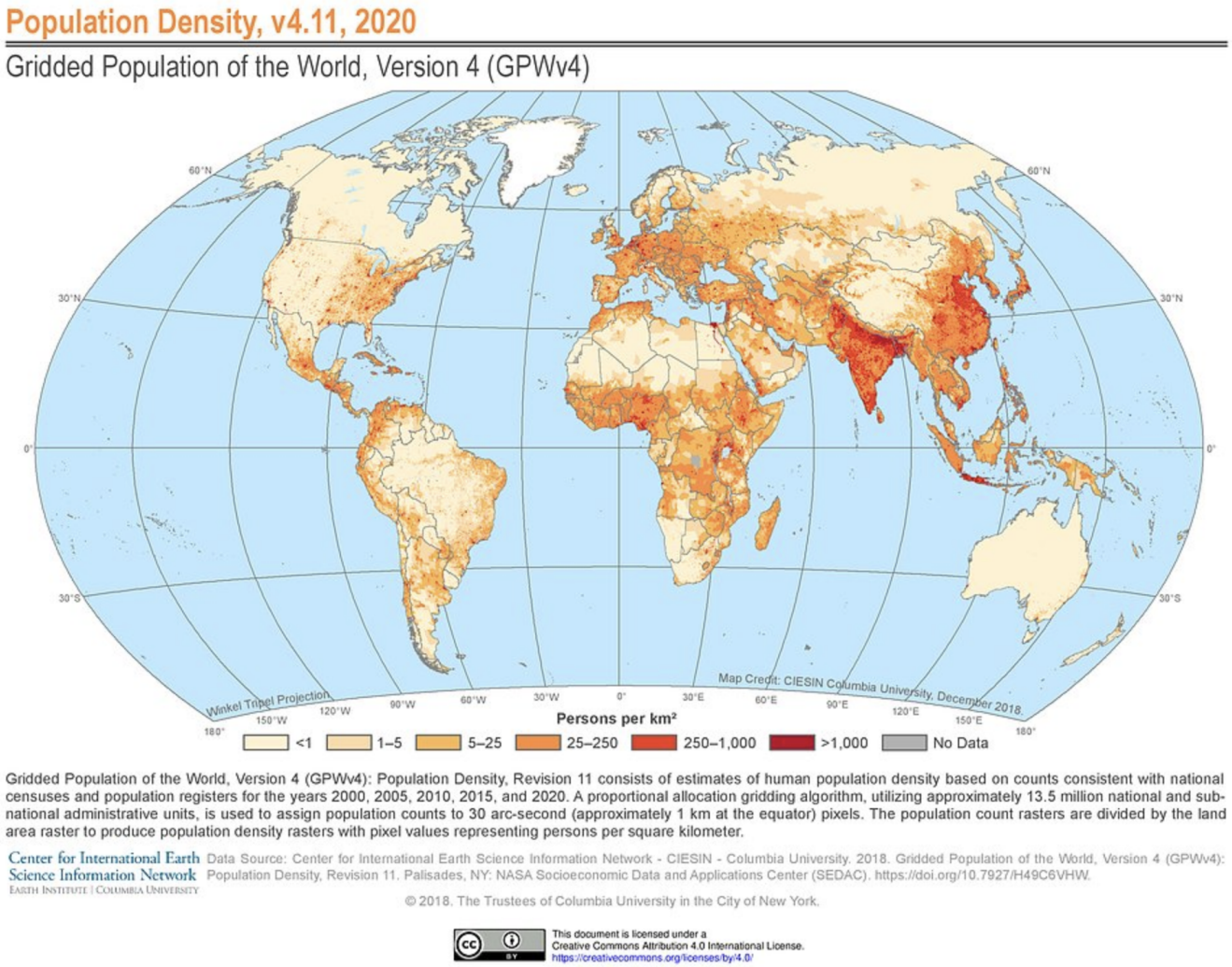


Figure 5.1.4.*a*: World density map. Graph by SEDAC (CC-BY-3.0) Gridded population of the world (version 4) illustrating population density for the year 2020. Density is indicated with darker or lighter shading using the following categories in persons per square kilometer (1, 1-5, 5-25, 25-250, 250-1000, 1000, and no data). North America has highest densities on the east coast of the United States, southern Mexico, Central American countries, and western and eastern coastlines of South America. All of Europe is high density. Africa has the highest densities in countries in the middle of the continent. In Asia, India, eastern China along with Asian islands have very high densities. (CC BY 4.0; The Trustees of Columbia University in the City of New York)

Questions

1. The data on the above graph is overlaid onto a world map. Why do you think this is a powerful way to display data?
2. What question(s) are the authors trying to answer with this graph and table?
3. Which areas of the globe have the highest density levels (darkest density shading)?

4. Are there any areas of the globe that surprise you (either high or low) with their density? Why?
5. How can policy makers/world leaders/health organizations use this information? Provide at least 3 ideas.
6. What is missing from this map? Hint: It's very large!

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 5.1.4: Data Dive- World Population Densities is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **12.4: Data Dive- World Population Densities** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

5.1.5: Review

Summary

After completing this chapter you should be able to...

- Discuss the history of human population growth.
- Define demographic transition.
- Explain how total fertility rate and age structure affect current and future population growth rate.
- Interpret age structure diagrams.
- Discuss the long-term implications of unchecked human population growth with respect to carrying capacity and ecological footprints.
- Describe efforts to regulate human population size.

The world's human population is growing at an exponential rate. Humans have increased the world's carrying capacity through migration, agriculture, medical advances, and communication. The age structure of a population allows us to predict population growth. Unchecked human population growth could have dire long-term effects on our environment.

Attribution

Modified by Melissa Ha from Human Population Growth from *General Biology* by OpenStax (licensed under CC-BY)

This page titled 5.1.5: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **12.5: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

5.2: Water Resources

Chapter Hook

The aqueducts of Rome are a huge feat of ancient achievement and ingenuity. Not only did they provide water for drinking and bathing but contributed to agriculture, hygiene (through sewers and plumbing), culture, and industry. These stone waterways took about 500 years to complete (312 BC to 226 AD) and contributed significantly to the success of Rome. Water is essential for survival, so other ancient civilizations found ways to move water as well; however, none even came close to being as successful or magnificent.



Figure 5.2.*a*: Roman aqueduct. Image by Pixabay (Public Domain)

https://bio.libretexts.org/@go/page/109175

5.2.1: Fresh Water Supply and the Water Cycle

5.2.2: Water Usage

5.2.3: Water Scarcity and Solutions

5.2.4: Data Dive- Aqueducts of Rome

5.2.5: Review

Attribution

Modified by Rachel Schleiger (CC-BY-NC).

This page titled 5.2: Water Resources is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

5.2.1: Fresh Water Supply and the Water Cycle

Water, air, and food are the most important natural resources to people. Humans can live only a few minutes without oxygen, less than a week without water, and about a month without food. Water also is essential for our oxygen and food supply. Plants breakdown water and use it to create oxygen during the process of photosynthesis.

Human babies are approximately 75% water, and adults are 60% water. Our brain is about 85% water, blood and kidneys are 83% water, muscles are 76% water, and even bones are 22% water. We constantly lose water by perspiration. In temperate climates we should drink about two quarts of water per day, and people in hot desert climates should drink up to 10 quarts of water per day. Loss of 15% of body-water usually causes death.

Water Reservoirs

Water is the only common substance that occurs naturally on Earth in three forms: solid, liquid and gas. The **hydrosphere** is the area of Earth where water movement and water storage occurs. **Water reservoirs** are the locations where water is stored. (Note that this term can also refer to artificial lakes created by dams.) Water is found as a liquid on the surface (rivers, lakes, oceans) and beneath the surface (groundwater), as ice (polar ice caps and glaciers), and as water vapor in the atmosphere. Figure 5.2.1.a illustrates the average time that an individual water molecule may spend in the Earth's major water reservoirs. **Residence time** is a measure of the average time an individual water molecule stays in a particular reservoir.

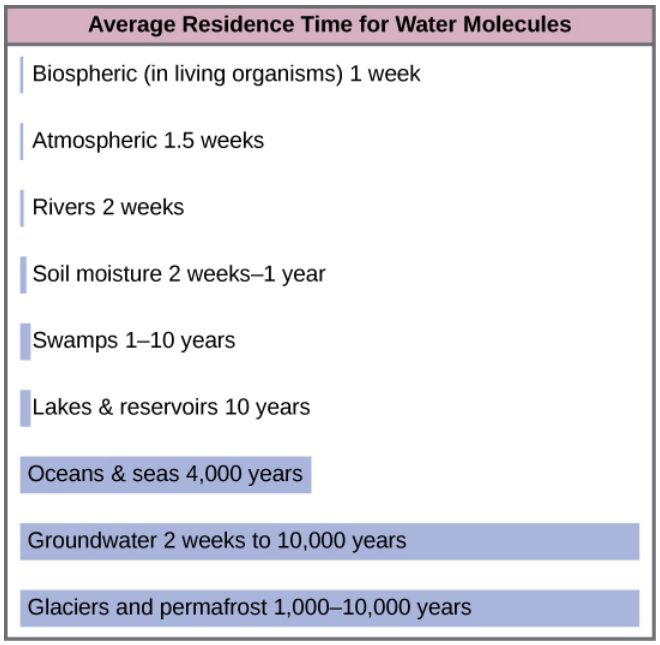
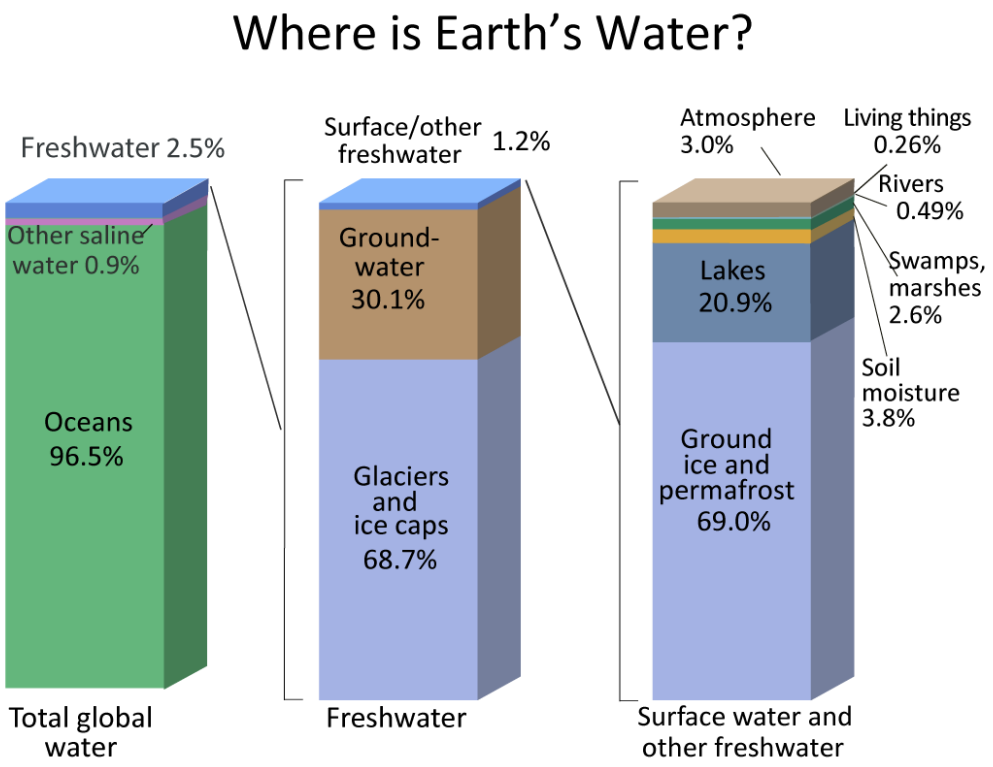


Figure 5.2.1.a: Average residence time that water remains in each reservoir. Water remains in organisms for about one week, in the atmosphere for 1.5 weeks, in rivers for two weeks, as soil moisture from two weeks to a year, in swamps for 1-10 years, in lakes for 10 years, in oceans and seas for 4,000 years, as groundwater for 2 weeks to 10,000 years, and in glaciers or as permafrost for 1,000-10,000 years. Image from OpenStax (CC-BY).

Earth is truly the Water Planet. The abundance of liquid water on Earth's surface distinguishes us from other bodies in the solar system. About 71% of Earth's surface is covered by water, and approximately half of Earth's surface is obscured by clouds (also made of water) at any time. There is a very large volume of water on our planet, about 1.4 billion cubic kilometers (km^3) = 330 million cubic miles, or about 53 billion gallons per person on Earth. All of Earth's water could cover the United States to a depth of 145 km (90 mi).

Despite the vast volumes of water on Earth, only 2.5% is fresh water (figure 5.2.1.b), and only 0.01% is available for humans to use. If all of world's water was shrunk to the size of 1 gallon, then the total amount of fresh water would be about 1/3 cup, and the amount of readily usable fresh water would be 2-3 tablespoons. Much of Earth's fresh water is trapped in glaciers and polar ice caps (figure 5.2.1.c), and this water is inconveniently located, mostly in Antarctica and Greenland. Shallow **groundwater** (water located below the Earth's surface) is the largest reservoir of usable fresh water. Many organisms are dependent on surface waters, such as

lakes and rivers, which comprise a small fraction of fresh water on Earth. A lack of these surface waters can have negative effects on ecosystems.



Source: Igor Shiklomanov's chapter "World fresh water resources" in Peter H. Gleick (editor), 1993, Water in Crisis: A Guide to the World's Fresh Water Resources. (Numbers are rounded).

Figure 5.2.1.*b*: Bar chart of the distribution of Earth's water, including total global water, fresh water, and surface water. Oceans are the largest of Earth's water reservoirs, and only 2.5% of Earth's water is fresh water. Most of this fresh water (68.7%) is in glaciers and ice caps, 30.1% is groundwater, and 1.2% is surface and other fresh water. Of this 1.2%, 3.0% is found in the atmosphere, 0.26% in living things, 0.49% in rivers, 2.6% in swamps and marshes, 3.8% as soil moisture, 20.9% in lakes, and 69.0% as ground ice and permafrost. Image from USGS/Igor Shiklomanov (public domain).



Figure 5.2.1.*c*: Mountain Glacier in Argentina. Glaciers are the largest reservoir of fresh water, but they are not used much as a water resource directly by society because of their distance from most people. Source: Luca Galuzzi – www.galuzzi.it

The Water Cycle

The **water cycle** (hydrologic cycle) shows the movement of water through different reservoirs, which include oceans, atmosphere, glaciers, groundwater, lakes, rivers, and organisms (figure 5.2.1.*d*). Solar energy, which warms the oceans and other surface waters, and gravity drive the motion of water in the water cycle. This leads to **evaporation** (liquid water to water vapor) of liquid surface water, **sublimation** (ice to water vapor) of frozen water, and **transpiration** (the loss of water from plants to the atmosphere). When water in the soil is taken up by plant roots, it moves through tubes in the plant (the vascular system), evaporates within the space of the leaf, and transpires through the stomata (small microscope openings) of the leaves. Ecologists combine transpiration and evaporation into a single term that describes water returned to the atmosphere: **evapotranspiration**. Thus, large amounts of water move into the atmosphere as water vapor.

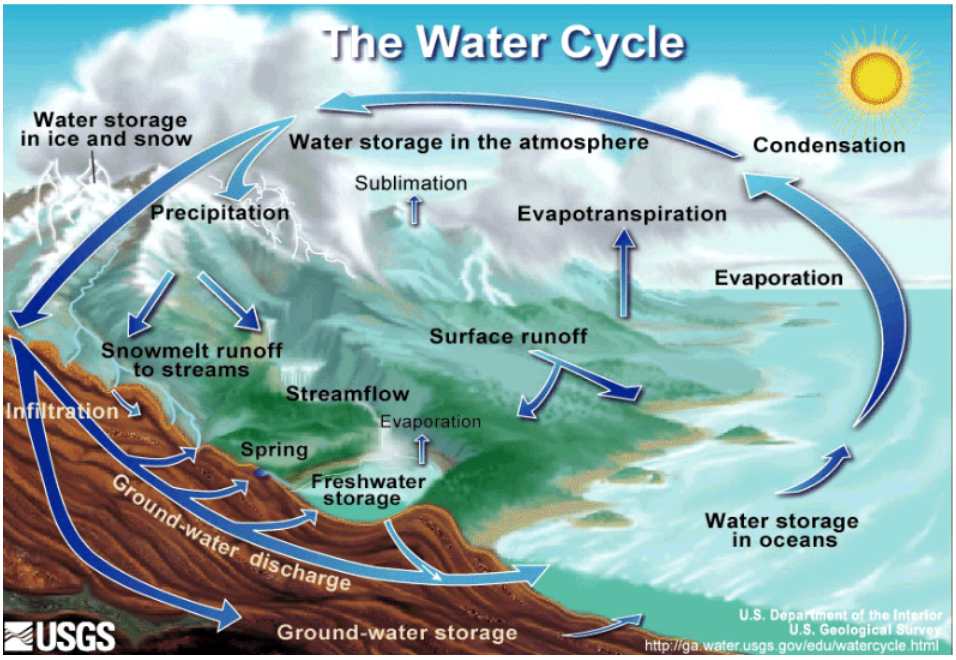


Figure 5.2.1.*d*: The water cycle. Arrows depict movement of water to different reservoirs located above, at, and below Earth's surface. Evapotranspiration and sublimination add water to the atmosphere, where it condenses to form clouds. Water is thus stored in the atmosphere. Through precipitation, water returns to the Earth's surface. When precipitation occurs at high elevation, water is stored as snow or ice. When the snow melts, the resultant runoff can add to streams. Water from precipitation can also infiltrate to replenish groundwater storage or form surface runoff. Groundwater discharge, water from streams, and surface runoff flows into bodies of fresh water and ocean. Source: United States Geological Survey

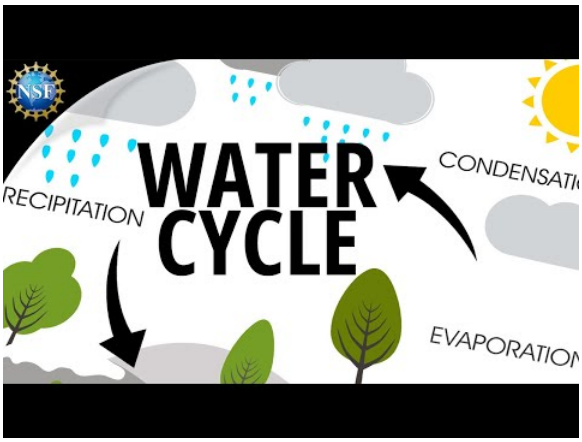
Water vapor in the atmosphere can migrate long-distances from ocean to over land by way of prevailing winds. Over the ocean or land, the air can cool and cause the water to **condense** back into liquid water. This usually happens in the form of very small water droplets that form around a microscopic piece of dust or salt called condensation nuclei. These small water droplets are visible as a cloud. Clouds build and once the water droplets are big enough, they fall to earth as **precipitation** (rain, snow, hail, or sleet), which returns water to Earth's surface.

Precipitation that reaches land can immediately return to the atmosphere, add to groundwater, or form surface runoff. In most natural terrestrial environments, rain encounters vegetation before it reaches the soil surface. A significant percentage of water evaporates immediately from the surfaces of plants or directly from the soil surface. Groundwater is replenished when water **infiltrates** into the soil and finally fills the pore spaces between particles in dirt, sand, and gravel or in the fissures in rocks. Groundwater slowly moves through rock and unconsolidated materials and some of it eventually reaches the surface again, where it discharges as springs and into streams, lakes, and the ocean. Many streams flow not because they are replenished from rainwater directly but because they receive a constant inflow from the groundwater below. Also, surface water in streams and lakes can infiltrate again to recharge groundwater. Therefore, the surface water and groundwater systems are connected. Groundwater can also ultimately flow into the ocean through subsurface groundwater flow, but some groundwater is found very deep in the bedrock and can persist there for millennia. **Surface runoff** is the flow of fresh water over land either from rain or melting ice. Runoff can

make its way through streams and lakes to the oceans. Surface runoff will occur only if the soil becomes saturated with water in a heavy rainfall.

The steps of the water cycle are explained in the video below.





Salinity and the Water Cycle

An important part of the water cycle is how water varies in salinity, which is the abundance of dissolved ions in water. The salt water in the oceans is highly saline, with about 35,000 mg of dissolved ions per liter of seawater. Evaporation is a distillation process that produces nearly pure water with almost no dissolved ions. As water vaporizes, it leaves the dissolved ions in the original liquid phase. Eventually, condensation forms clouds and sometimes precipitation. After rainwater falls onto land, it dissolves minerals in rock and soil, which increases its salinity. Rain and surface runoff are major ways in which minerals, including phosphorus and sulfur, are cycled from land to water. The environmental effects of runoff were discussed in Biogeochemical Cycles. Fresh water (such as lakes, rivers, and near-surface groundwater) has a relatively low salinity.

Human Interactions with The Water Cycle

Humans alter the water cycle by extracting large amounts of fresh water from surface waters as well as groundwater (see Water Usage). Additionally, changes in land use, such as deforestation, agriculture, and urbanization, reduce vegetation cover, which reduces infiltration and increases surface runoff. (Vegetation naturally traps precipitation as it falls, slows the flow surface runoff, and increases the infiltration rate.) This intensifies flooding and exacerbates erosion, lowering soil quality and causing sediment pollution in water. Additionally, humans redirect the flow of water by building dams and aqueducts (figure $5.2.1.e$). So much water

is removed or redirected from the Colorado River in the western United States that, despite its considerable size, in some years it is dry before reaching the sea in Mexico. In an extreme example, the Aral Sea in Central Asia has decreased to only 10% of its initial size after water was diverted for agriculture (see this case study for more details).



Figure 5.2.1.*e*: The California Aqueduct carries water needed for agriculture from Northern California to Southern California. Image by USGS (public domain).

Water Resources

Freshwater resources are ultimately replenished by precipitation. This water can then be obtained from surface water, such as rivers and lakes, and from aquifers, which store groundwater.

Primary Fresh Water Resources: Precipitation

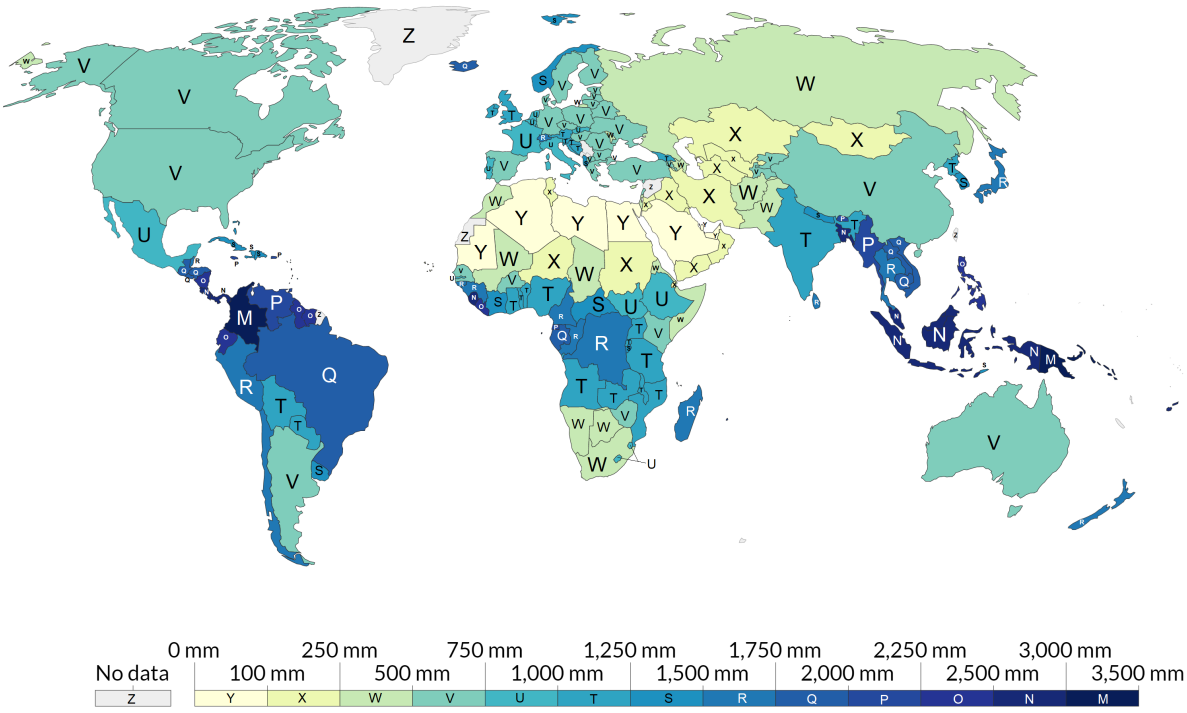
Precipitation levels are unevenly distributed around the globe, affecting fresh water availability (figure 5.2.1.*f*). In general, due the uneven heating of the Earth and global air circulation cells resulting from the Earth's rotation, air rises near the equator and near 60° north and south latitude and sinks at the poles and 30° north and south latitude. As discussed in Climate Effects on Biomes, the intense sunlight at the equator heats air, causing it to rise and cool, which decreases the ability of the air mass to hold water vapor and results in frequent rainstorms. Around 30 degrees north and south latitude, descending air conditions produce warmer air, which increases its ability to hold water vapor and results in dry conditions. Both the dry air conditions and the warm temperatures of these latitude belts favor evaporation.

The size of continents, mountains, prevailing winds, ocean circulation patterns, and even the proximity of bodies of water can affect local climate patterns. For example, when cold winds blow across the relatively warm Great Salt Lake, the air warms, which causes it to pick up moisture. This local increase in the moisture content of the air may eventually fall as snow or rain on nearby mountains, a phenomenon known as “lake-effect precipitation”.

Average annual precipitation, 2014

Precipitation in millimeters per year. Precipitation is defined as any kind of water that falls from clouds as a liquid or a solid.

Our World in Data



Source: World Bank

CC BY

Figure 5.2.1.*f*: Annual average precipitation by country in 2014. Countries with the highest rainfall (3000-3500 mm) are Nicaragua, Costa Rica, Panama, Columbia, Sierra Leone (in West Africa), Bangladesh, and many of the islands is Southeast Asia. Countries with the least rainfall (0-100 mm) are Saudi Arabia and several North African countries. The United States has intermediate rainfall (500-750 mm). The world's rainforests grow in areas with high rainfall while areas with very little rainfall are deserts. Image modified from Our World in Data/World Bank (CC-BY).

In the United States, the 100th Meridian roughly marks the boundary between the humid and arid parts of the country (figure 5.2.1.*g*). Irrigation is required to grow crops west of the 100th Meridian. In the West, surface water is stored in reservoirs (artificial lakes) and mountain snowpacks and strategically released through a system of canals during times of high use.

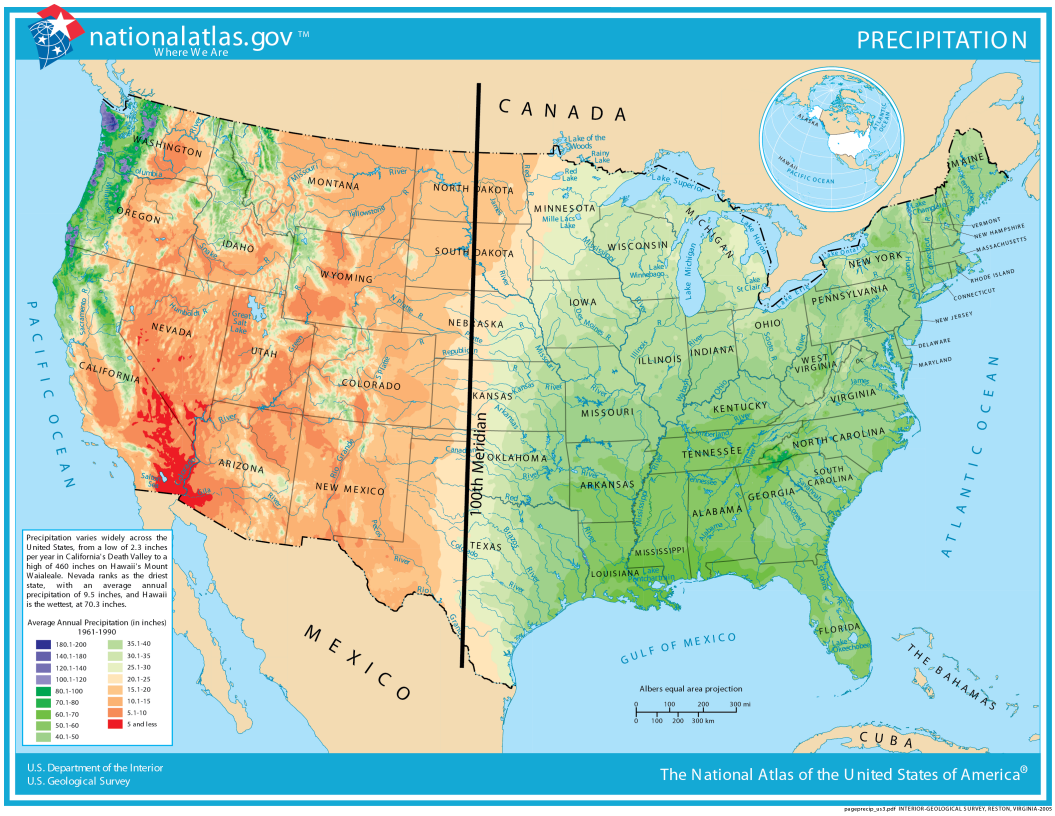


Figure 5.2.1.g: Distribution of precipitation in the United States. The 100th Meridian is approximately where the average precipitation transitions from relatively wet to dry. Eastern states are mostly green, indicating greater precipitation than most western states, which are mainly orange or red (indicating low precipitation). (Source: U.S. Geological Survey)

Surface Water: Rivers and Lakes

Rivers are an important water resource for irrigation of cropland and drinking water for many cities around the world. Flowing water from rain and melted snow on land enters river channels by surface runoff (figure 5.2.1.h) and seepage from surrounding land. The geographic area drained by a river and its tributaries is called a **watershed**. The Mississippi River watershed includes approximately 40% of the U.S., a measure that includes the smaller watersheds, such as the Ohio River and Missouri River that help to comprise it. Rivers that have had international disputes over water supply include the Colorado (Mexico, southwest U.S.), Nile (Egypt, Ethiopia, Sudan), Euphrates (Iraq, Syria, Turkey), Ganges (Bangladesh, India), and Jordan (Israel, Jordan, Syria).

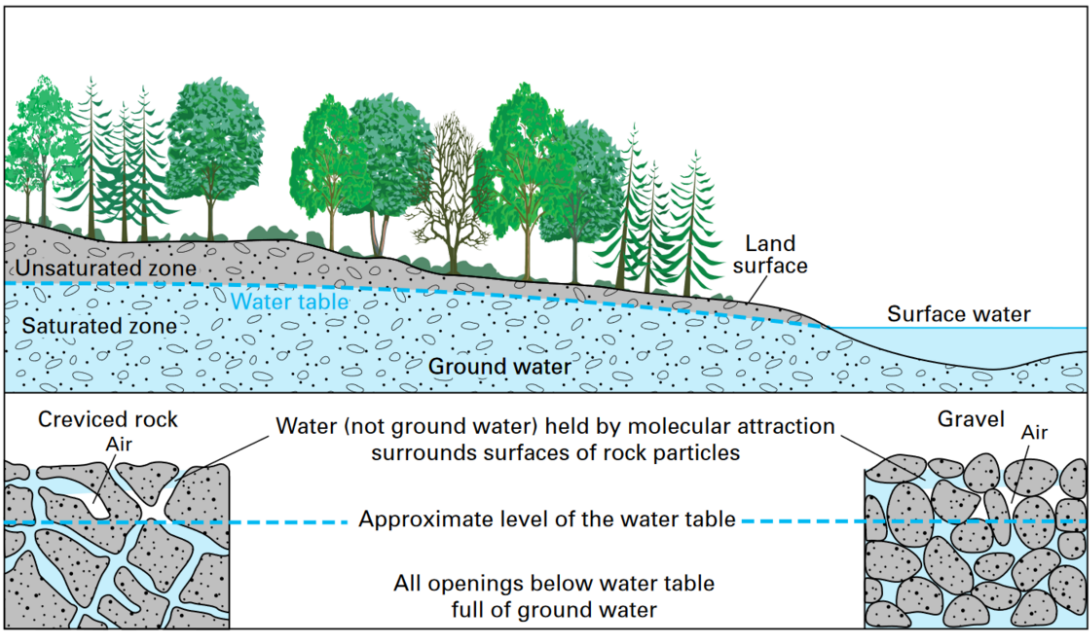


Figure 5.2.1.*h*: Surface runoff, part of overland flow in the water cycle Source: James M. Pease at Wikimedia Commons

In addition to rivers, lakes can also be an excellent source of fresh water for human use. They usually receive water from surface runoff and groundwater. By building dams, people create artificial lakes (**reservoirs**).

Groundwater Resources

Although most people in the world use surface water, groundwater is a much larger reservoir of usable fresh water, containing more than 30 times more water than rivers and lakes combined. A large area of subsurface, porous rock unit or sediment that holds extractable groundwater is an **aquifer**. The **saturated zone** of an aquifer is where groundwater completely fills pore spaces in earth materials. The **water table** is the uppermost level at which the pores are fully saturated with water (figure 5.2.1.*i*).



How ground water occurs in rocks.

Figure 5.2.1.*i*: The unsaturated and saturated zones of an aquifer, which stores groundwater. The water table is the upper level of the saturated zone. The land surface is covered in vegetation, and the top of the surface water is continuous with the water table. The bottom of the image shows creviced rock (left) and gravel (right), both of which contain spaces that store air and water. Above the water table, water (not groundwater) is held by molecular attraction and surrounds the surfaces of rock particles. All openings below the water table are full of water table. Image by USGS (public domain)

A combination of a place to put water (porosity) and the ability to move water (permeability) makes a good aquifer. **Porosity** is a measure of the open space in rocks –expressed as the percentage of open space that makes up the total volume of the rock or sediment material. **Permeability** is a measure of the interconnectedness of pores in a rock or sediment. The connections between pores allows for that material to transmit water. Porosity and permeability are functions of soil particle composition. For example, clays generally have very high porosity, but the pores are poorly connected, thereby causing low permeability.

Aquifers are commonly drilled, and wells installed, to provide water for agriculture and personal use. In many cases aquifers are being depleted faster than they are being replenished by water infiltrating down from above. Groundwater is a particularly important resource in arid climates, where surface water may be scarce. In addition, groundwater is the primary water source for rural homeowners, providing 98% of that water demand in the U.S.

As groundwater is pumped from water wells, there usually is a localized drop in the water table around the well called a **cone of depression** (figure 5.2.1.*j*). When there are a large number of wells that have been pumping water for a long time, the regional water table can drop significantly. This is called **groundwater mining**, which can force the drilling of deeper, more expensive wells that commonly encounter more saline groundwater. Rivers, lakes, and artificial lakes (reservoirs) can also be depleted due to overuse. Some large rivers, such as the Colorado in the U.S. and Yellow in China, run dry in some years.

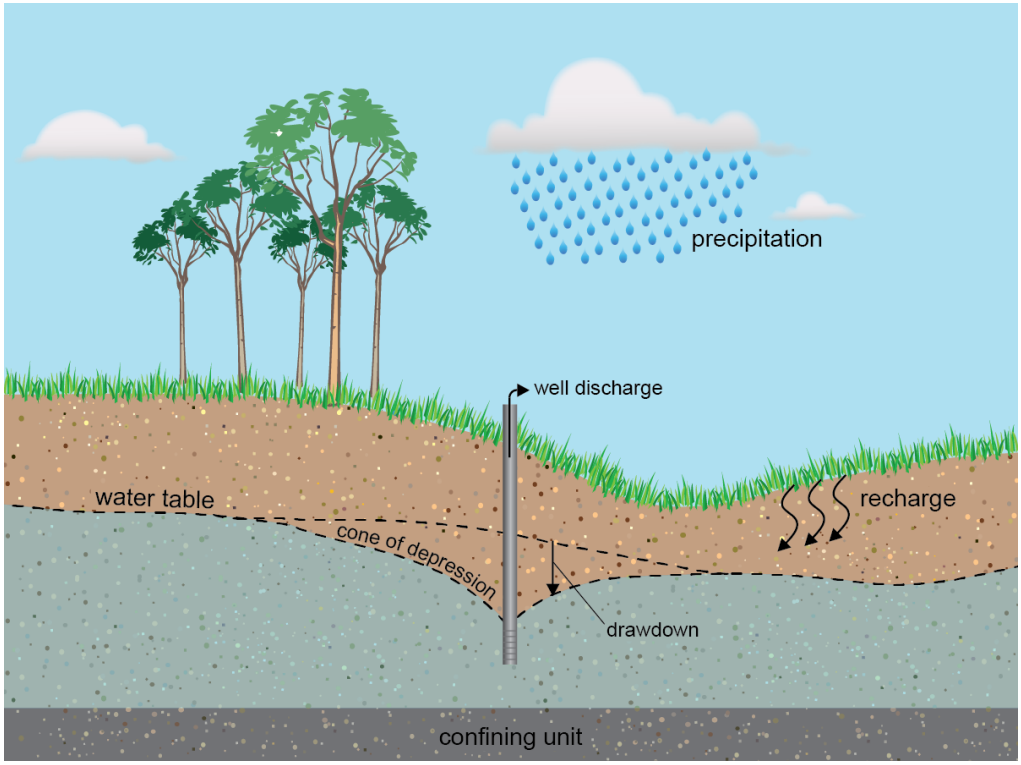


Figure 5.2.1.j: As groundwater is extracted from the well, the water table of the aquifer dips, forming a cone of depression. Drawdown is between the water table and the cone of depression. Precipitation is represented by water droplets falling from a cloud. It soaks into the ground, recharging the aquifer. At the bottom of the aquifer is the confining unit, which is impermeable to water. Image by Tara Gross/USGS (public domain).

Another water resource problem associated with groundwater mining is **saltwater intrusion**, where overpumping of freshwater aquifers near ocean coastlines causes salt water to enter freshwater zones. The drop of the water table around a cone of depression in an aquifer can change the direction of regional groundwater flow, which could send nearby pollution toward the pumping well instead of away from it. Finally, problems of **subsidence** (gradual sinking of the land surface over a large area) and **sinkholes** (rapid sinking of the land surface over a small area) can develop due to a drop in the water table. Because the pores in the aquifer collapse as subsidence occurs, this permanently reduces the capacity of the aquifer to hold water in the future.

Groundwater is replenished through infiltration, seepage from surface water (lakes, rivers, reservoirs, and swamps), surface water deliberately pumped into the ground, irrigation, and underground wastewater treatment systems (septic tanks). **Recharge areas** are locations where surface water infiltrates into the ground rather than running into rivers or evaporating (figure 5.2.1.k). Recharge areas are generally the topographically highest location of an aquifer. They are characterized by streams that lie below the water table and sediment or rock that allows infiltration into the subsurface. Wetlands, for example, are excellent recharge areas. Recharge areas mark the beginning of groundwater flow paths.

Recharge can be induced through the aquifer management practice of aquifer storage and recovery. Injection wells allow for humans to increase the rate of recharge into an aquifer system by pumping water into an aquifer (figure 5.2.1.k). Injection wells are regulated by state and federal governments to ensure that the injected water is not negatively impacting the quality or supply of the existing groundwater in the aquifer. Some aquifers are capable of storing significant quantities of water, allowing water managers to use the aquifer system like a surface reservoir. Water is stored in the aquifer during periods of low water demand and high water supply and later extracted during times of high water demand and low water supply.

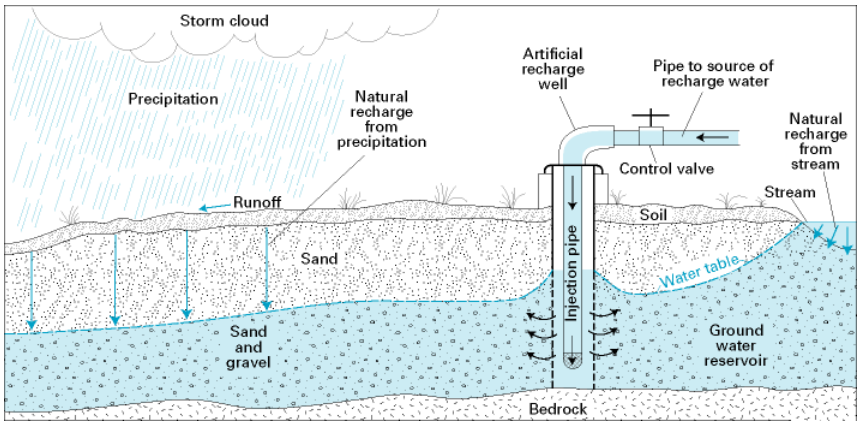


Figure 5.2.1.*k*: Groundwater can be replenished through natural recharge areas as well as artificial injection wells. In this diagram, precipitation emerges from a storm cloud. Some of the water washes away as runoff, and some water infiltrates through the sand (natural recharge from precipitation). The water fills the sand and gravel in the aquifer (labeled ground water reservoir). Below the aquifer is bedrock, which is impermeable to water. Another natural source of recharge is from streams. Through an artificial recharge well, water flows from a pipe through a control valve into an injection pipe. This also adds to the groundwater. Image by USGS (public domain).

Attributions

- Water Cycle and Fresh Water Supply, Water Supply Problems and Solutions, and Biogeochemical Cycles from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- The Water Cycle, Water Use and Distribution, and Groundwater from *An Introduction to Geology* by Chris Johnson et al. (licensed under CC-BY-NC-SA)
- Biogeochemical Cycles and the Flow of Energy in the Earth System from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors (licensed under CC-BY). Download for free at CNX.
- Biogeochemical Cycles from *Biology 2e* by OpenStax (CC-BY). Access for free at openstax.org.

This page titled 5.2.1: Fresh Water Supply and the Water Cycle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **13.1: Fresh Water Supply and the Water Cycle** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

5.2.2: Water Usage

Freshwater supply is one of the most important ecosystem services. In 2014, global water consumption was 3999 km^3 per year (over 1,000 trillion gallons!). The greatest use of this water is for irrigation in agriculture, but significant quantities of water are also extracted for public and municipal use, as well as industrial applications and power generation (figure 5.2.2.a).

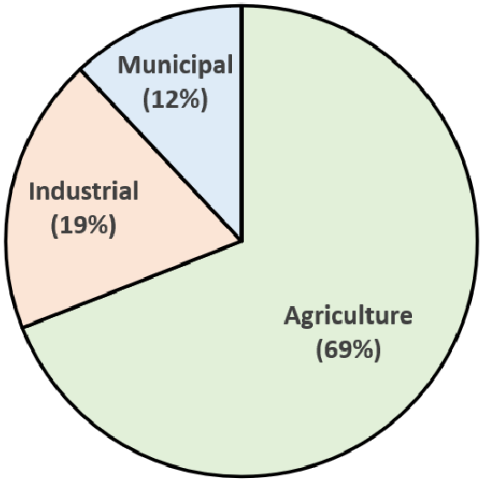


Figure 5.2.2.a: Global water usage by sector. Municipal refers to domestic use, such as by local businesses and households. Data from FAO. Image by Melissa Ha (CC-BY-NC).

Humans require only about 1 gallon per day to survive, but a typical person in a U.S. household uses approximately 80-100 gallons per day, which includes cooking, washing dishes and clothes, flushing the toilet, and bathing. Additionally, we rely on the food, energy, and mineral resources, which all require water to produce. For example, approximately three gallons are needed to produce a tomato, 150 gallons for a loaf of bread, and 1,600 for a pound of beef. (You can learn more about the water footprint of different foods using this interactive website.) Twenty-one gallons are needed to produce one kilowatt-hour (kWh) of electricity from a traditional power plant (about 1 kWh is needed to heat an oven for 30 minutes), and one ton of steel consumes about 63,000 gallons of water. The **water demand** of an area is a function of the population and other uses of water.

In the United States, 281 billion gallons of water were withdrawn per day in 2015, of which 82 billion gallons are fresh groundwater (figure 5.2.2.b). The state of California accounts for 9% of national water withdrawal (figure 5.2.2.c).

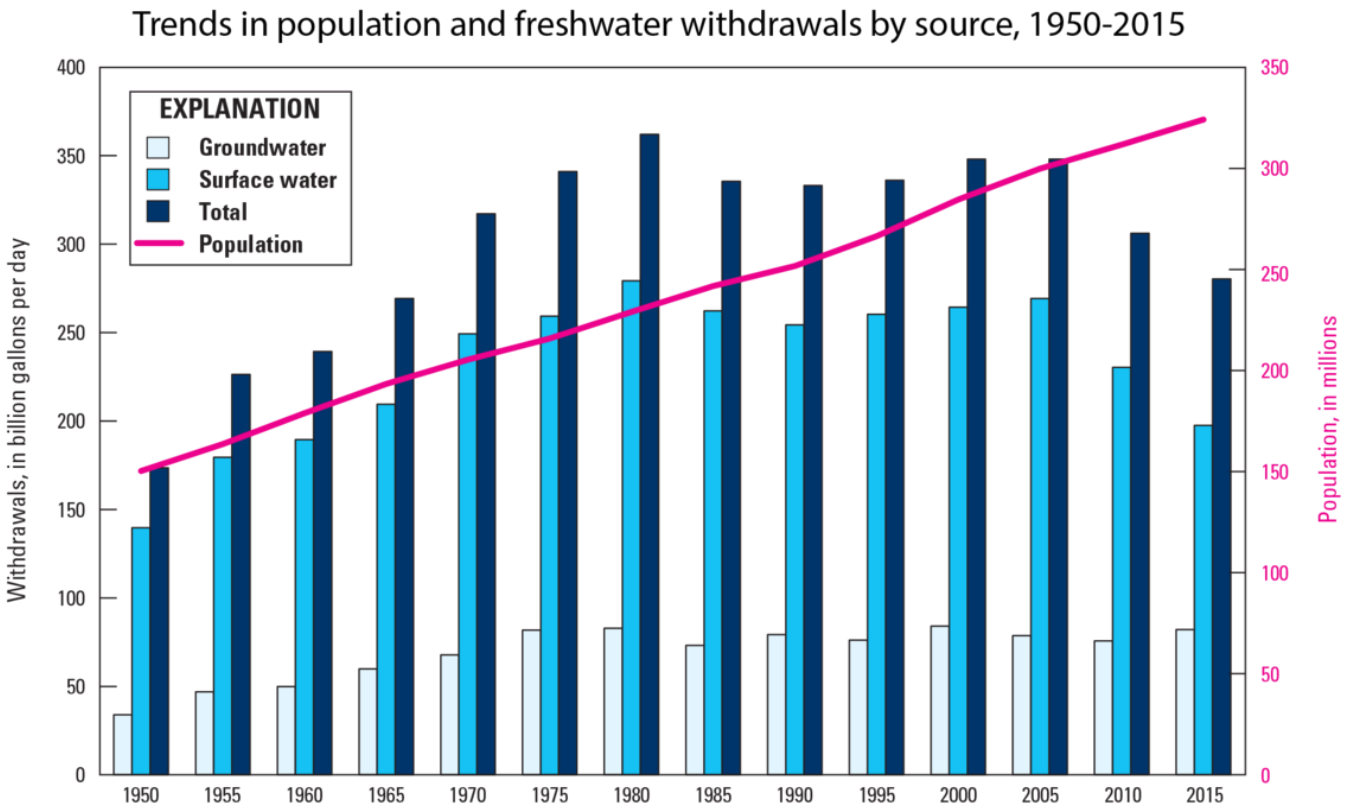


Figure 5.2.2.*b*: Average daily freshwater withdrawals (in billion gallons per day, left y-axis) over time (x-axis) in the United States. Light blue bars (left) represent groundwater withdrawals, medium blue (middle) represents withdrawals from surface water, and dark blue bars (right) represent total withdrawals in billions of gallons per day. In 1950, total withdrawals were around 170 billion gallons per day. This gradually increased until 1980 at just over 350 billion gallons per day. In 1985, total withdrawals dropped to around 325 billion gallons and stayed relatively constant until decreasing in 2010. In 2015, total withdrawals were around 260 billion gallons. Most of the changes in withdrawals overtime were due to changes in surface water withdrawals rather than groundwater withdrawals. Surface water withdrawals have generally been 2-3 times greater than groundwater withdrawals. The pink line represents the United States population size in millions (right y-axis), which has increased from roughly 170 million in 1950 to about 325 million in 2015. Image by USGS (public domain).

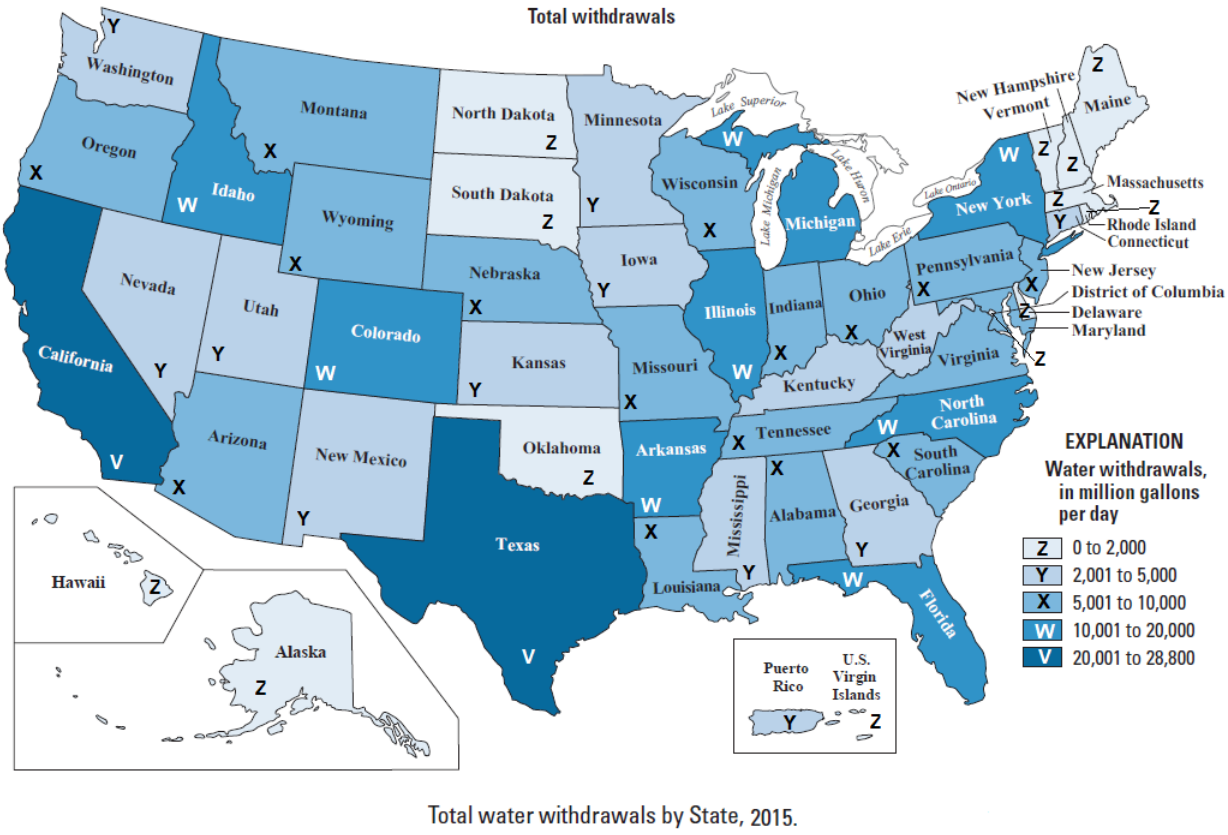


Figure 5.2.2.*c*: Total water withdrawals by state in 2015. The lightest blue (Z) indicates 0-2,000 million gallons per day, and the darkest blue (V) indicates 20,001-28,800 million gallons per day. California and Texas consume the most water, and Alaska, Hawaii, the Dakotas, Oklahoma, and some New England states consume the least. Image modified from USGS (public domain).

Attributions

Modified by Melissa Ha from the following sources

- Water Use and Distribution from *An Introduction to Geology* by Chris Johnson et al. (licensed under CC BY-NC-SA)
- Water Cycle and Fresh Water Supply from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Biogeochemical Cycles and the Flow of Energy in the Earth System from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors (licensed under CC-BY). Download for free at CNX.

This page titled 5.2.2: Water Usage is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **13.2: Water Usage** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

5.2.3: Water Scarcity and Solutions

One of the most important environmental goals is to provide clean water to all people. Fortunately, water is a renewable resource and is difficult to destroy. Evaporation and precipitation combine to replenish our fresh water supply constantly; however, water availability is complicated by its uneven distribution over the Earth.

Water Scarcity

The **water crisis** refers to a global situation where people in many areas lack access to sufficient water, clean water, or both. Arid climate and densely populated areas have combined in many parts of the world to create water shortages, which are projected to worsen in the coming years due to population growth, water overuse, water pollution, and climate change. Specifically, climate change shifts precipitation patterns and causes the snow pack that recharges rivers to melt earlier in the year. Furthermore, rising sea levels associated with climate change worsen saltwater intrusion.

Water scarcity refers to water shortages, which can be physical or economic (figure 5.2.3.*a*). **Physical water scarcity** is the lack of sufficient water resources in an area; that is, water is depleted more quickly than it is replenished. Unpredictable precipitation patterns associated with climate change, which increase the risk of flooding and drought, exacerbates physical water scarcity. **Economic water scarcity** occurs when people cannot afford access to water. The United Nations estimates that over half of the global population faces water scarcity for one or more months of the year (see The Sustainable Development Goals Report 2019). According to the World Health Organization and UNICEF, 785 million people lack access to even a basic drinking water service (see Drinking Water) and two billion people lack access to improved sanitation as simple as a pit latrine (see Sanitation), and three billion people lack a facility to wash their hands (see Hand Hygiene for All). As a result, nearly 829,000 people die every year from diarrheal diseases, and 297,000 of those deaths occur among children under the age of five (see Drinking Water).

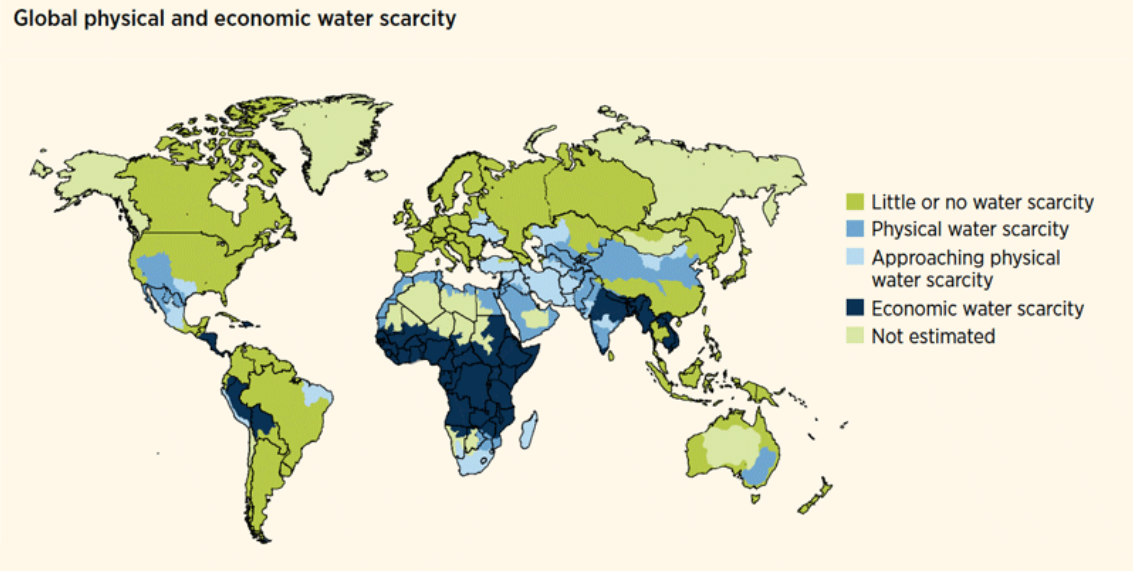


Figure 5.2.3.*a*: Regions facing or approaching water scarcity in 2012. Physical water stress occurs in the southwestern United States, the Middle East, North Africa, southern India, and southeast Australia. Economic water stress occurs in Central and East Africa, western India, and Southeast Asia. Image by World Water Assessment Programme (WWAP; public domain).

Solutions for Addressing Water Shortages

While some human activities have exacerbated the water crisis, humans have also developed technologies to better acquire or conserve freshwater. Solutions to addressing water shortages include dams and reservoirs, rainwater harvesting, aqueducts, desalination, water reuse, and water conservation.

Dams and Reservoirs

Reservoirs (artificial lakes) that form behind dams in rivers can collect water during wet times and store it for use during dry spells (figure 5.2.3.*b*). They also can be used for urban water supplies. Other benefits of dams and reservoirs are hydroelectricity, flood

control, and recreation. Some of the drawbacks are evaporative loss of water in arid climates and downstream river channel erosion. Additionally, dams reduce water flow downstream, which could lead to political conflicts when rivers span states or countries.

The negative ecosystem impacts of dams are another major drawback. For example, dams change a river to a lake habitat and interfere with migration and spawning of fish. Furthermore, warming of the surface water in the reservoir influences the temperature of the water downstream, impacting the fish and aquatic invertebrates that are adapted to colder water. Dams also trap sediments that would otherwise continue to flow down the river, creating habitat and supplying nutrients downstream.



Figure 5.2.3.b: The Hoover Dam on the Colorado River in Nevada. Behind the dam is Lake Mead, the largest reservoir in the United States. Image by Dsimic (CC-BY-SA).

Rainwater Harvesting

Rainwater harvesting involves catching and storing rainwater before it reaches the ground. Figure 5.2.3.c shows a complex rainwater harvesting system (rain water capture system) proposed for federal buildings, but smaller, simpler systems (sometimes called rain barrels) can be used by individual homeowners (figure 5.2.3.d).

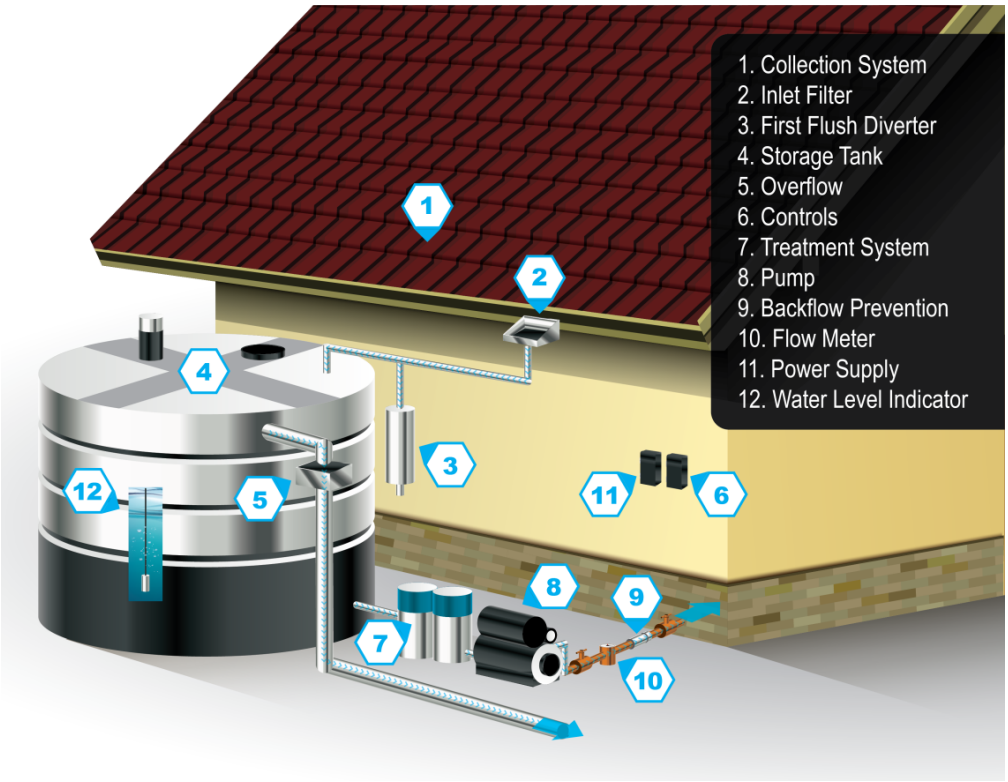


Figure 5.2.3.*c*: The components of a rainwater harvesting system include (1) the collection system consisting of the roof surface and gutters that capture rain water and send it to (2) the inlet filter, a screen that catches large debris. (3) The first flush diverter removes debris not captured by the inlet filter from the initial stream of rainwater. (4) The storage tank is composed of food-grade polyester resin material approved by the U.S. Food and Drug Administration (FDA), which is green in color and helps to reduce bacterial growth. (5) The overflow is a drainage spout that releases excess water if the storage tank is full. The (6) control system monitors the water level and filtration system, and the (7) treatment system is responsible for filtration and disinfection, treating water until it is potable (safe to drink). The (8) pump moves water through the system to where it will be used, and the (9) backflow preventer ensures that water cannot flow backwards through the system in cases of low water pressure (when the storage tank is low). (10) The flow meter (with data logger) measure water production. The (11) power supply may rely on conventional energy sources or solar energy, and the (12) water level indicator monitors the water level in the storage tank. Caption modified from and image from Office of Energy Efficiency and Renewable Energy, Department of Energy (public domain).



Figure 5.2.3.*d*: A rain barrel is a simple rainwater harvesting system. Homeowners commonly use the stored water to irrigate their gardens. Image by Walton LaVonda, USFWS (public domain).

Aqueducts

Aqueducts can move water from where it is plentiful to where it is needed. Aqueducts can be controversial and politically difficult especially if the water transfer distances are large. One drawback is the water diversion can cause drought in the area from where the water is drawn. For example, Owens Lake and Mono Lake in central California began to disappear after their river flow was diverted to the Los Angeles aqueduct (figure 5.2.3.*e*). Without water supply, Owens Lake dried and became a major source of particulate matter, polluting the air during dust storms (see Air Pollution). Owens Lake remains almost completely dry, but Mono Lake has recovered more significantly due to legal intervention. Learn more about the Los Angeles Aqueduct here.



Figure 5.2.3.*e*: The Los Angeles Aqueduct diverts the flow of several eastern California rivers to the city of Los Angeles. Image by Los Angeles (CC-BY-SA).

Desalination

One method that can actually increase the amount of freshwater on Earth is **desalination**, which involves removing dissolved salt and minerals from seawater or saline groundwater (figure 5.2.3.*f*). An advantage of this approach is that there is a virtually unlimited supply of saltwater. There are several ways to desalinate seawater including boiling, filtration, electrodialysis (applying an electric current to removed the ions which comprise salts), and reverse osmosis (figure 5.2.3.*g*). All of these procedures are moderately to very expensive and require considerable energy input, making the water produced much more expensive than freshwater from conventional sources. In addition, the process creates highly saline wastewater, which must be disposed of and creates significant environmental impact. Desalination is most common in the Middle East, where energy from oil is abundant but water is scarce.



Figure 5.2.3.f: Port Stanvac Desalination Plant under construction in South Australia. Image by Vmenkov (CC-BY-SA).

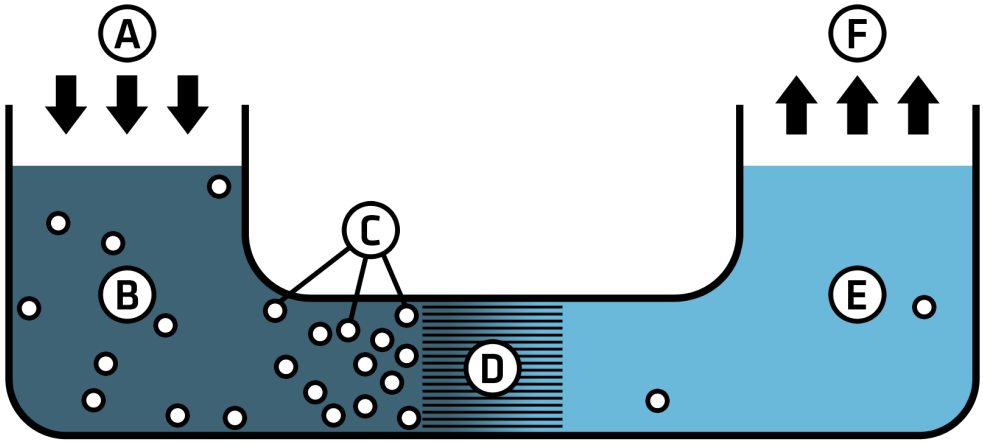


Figure 5.2.3.g: Reverse osmosis is one method of desalination. In this process, pressure (A) is applied to salt water, which contains ions (charged atoms or molecules) and other contaminants (C). The water is forced through a semipermeable membrane (D), through which the contaminants cannot pass. Clean freshwater that is safe to drink (E) is then distributed (F). Image by Colby Fisher (CC-BY-SA).

Water Reuse (Water Recycling)

Water recycling refers to reusing water for appropriate purposes such as agriculture, municipal water supply, industrial processes, and environmental restoration (figure 5.2.3.h). This could occur at the scale of a single household, for example, installing plumbing that reroutes water drained from the sink to flush the toilet. Water recycling can also occur at large scales. For example, wastewater from the sewage system is regularly treated to an extent, but it can be treated further to produce **potable water** (which is safe to drink) and then pumped into depleted aquifers. This approach limits saltwater intrusion of aquifers near the coast and reduces dependence on precipitation and subsequent infiltration to recharge aquifers. Orange County Water District in California employed this system following an information campaign to explain the purification process and ensure public confidence in the safety of the treated wastewater.

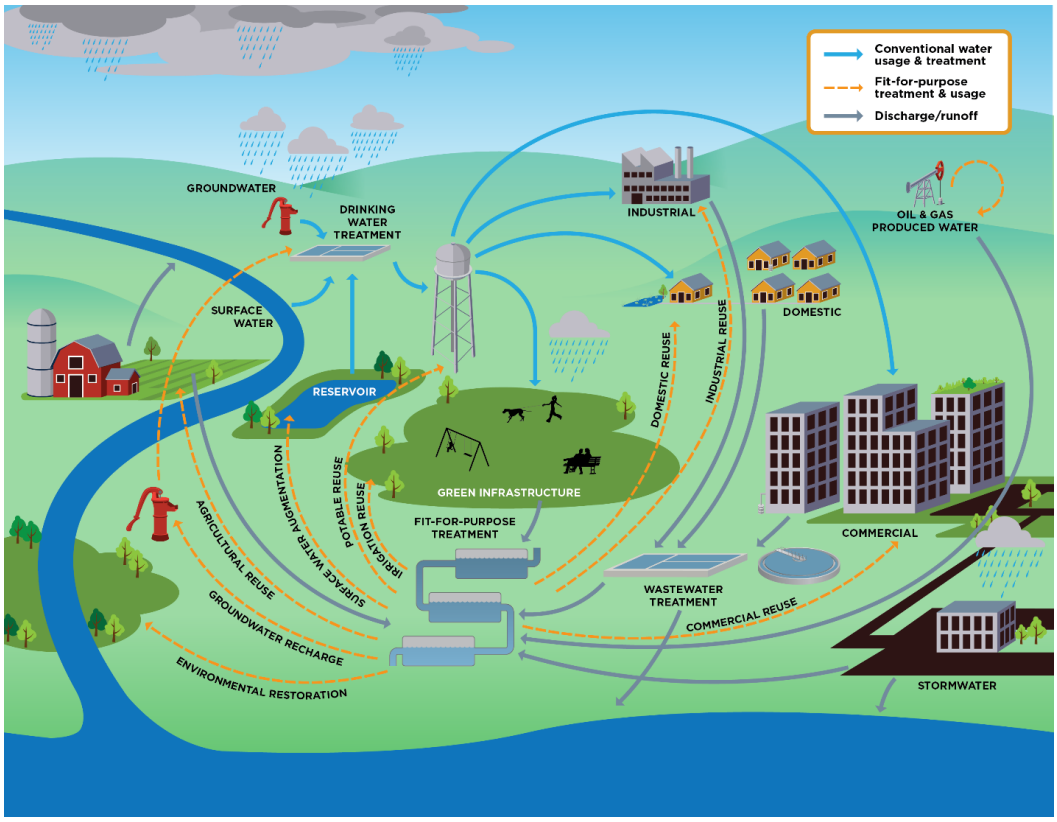


Figure 5.2.3.*h*: Flowchart of water reuse. Surface water, groundwater, and water from reservoirs undergoes drinking water treatment. From there, it serves commercial, industrial, and domestic purposes and supports green infrastructure. Used water is treated to the extent needed for its purpose (fit-for-purpose treatment). Fit-for-purpose treatment and usage include environmental restoration, ground water recharge, agricultural reuse, surface water augmentation, portable reuse, irrigation reuse, domestic reuse, industrial reuse, and commercial reuse. Discharge from industrial, domestic, and commercial buildings is processed through wastewater treatment. Image by EPA (public domain).

Water Conservation

Water conservation refers to using less water and using it more efficiently. Around the home, conservation can involve both water-saving technologies and behavioral decisions. Examples of water-saving technologies include high-efficiency clothes washers and low-flow showers and toilets. Water-conserving behaviors include turning off the water while you brush your teeth, taking shorter showers and showers instead of baths, and fixing leaky faucets. A dishwasher uses less water than washing dishes by hand, particularly the dishwasher is only run when it is full. Similarly, running fewer, larger loads of laundry conserves water relative to more frequent, smaller loads. Choosing foods with a low water footprint (like eggs) over those with a high water footprint (like beef) can also conserve water.

Gardening offers several water-saving opportunities. If you live in a dry climate, consider growing only native, drought-tolerant vegetation, which requires little irrigation (figure 5.2.3.*h*). When you do irrigate your garden, do so only as needed and early in the morning, when less water will be lost to evaporation. Drip systems assist in delivering only the needed amount of water in a way that minimizes evaporation. These strategies can also be applied at large scales in agriculture, which is extremely important considering the high agricultural demands on our water supply relative to municipal use. Water conservation strategies in agriculture include growing crops in areas where the natural rainfall can support them, more efficient irrigation systems such as drip systems, and no-till farming, which reduces evaporative losses by covering the soil.



Figure 5.2.3.7: Using water-efficient plants and irrigation not only looks beautiful, but saves resources and supports biodiversity and natural pollinators. Image and caption from CA Department of Water Resources (public domain).

Bottled water is not a sustainable solution to the water crisis. Bottled water is not necessarily any safer than the U.S. public water supply, it costs on average about 700 times more than U.S. tap water, and every year it uses approximately 200 billion plastic and glass bottles that have a relatively low rate of recycling. Compared to tap water, it uses much more energy, mainly in bottle manufacturing and long-distance transportation. (Purchasing a water filter is a more sustainable solution than bottled water if you do not like the taste of tap water.)

References

Drinking Water. 2019. WHO. Accessed 2020-12-29.

Hand Hygeine for All. 2020. UNICEF. Accessed 2020-12-29.

Sanitation. 2019. WHO. Accessed 2020-12-29.

The Sustainable Development Goals Report. 2019. United Nations. Accessed 2020-12-29.

Attribution

Modified by Melissa Ha from the following sources:

- Water Availability and Use from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Basic Information about Water Reuse by United States Environmental Protection Agency (public domain)

This page titled 5.2.3: Water Scarcity and Solutions is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **13.3: Water Scarcity and Solutions** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

5.2.4: Data Dive- Aqueducts of Rome

Overview

All drinking water sources are subject to contamination that require appropriate treatment to eliminate disease-causing pathogens. Sources of water contamination can originate from naturally occurring chemicals and minerals (arsenic, radon, uranium), local land use practices (fertilizers, pesticides), manufacturing processes, and sewage. If left untreated, the presence of contaminants in water can lead to adverse health effects, both acute and chronic. This is why the achievement of the aqueducts of Rome are so astounding! Not only did they water to the residents of Rome, but somehow the Romans were able to identify the cleanest water sources for drinking without the most basic testing equipment. In 1995 Peter Aicher published a book titled, “Guide to the aqueducts of ancient Rome.” The book discusses a lot of the history surrounding the 11 aqueducts while also integrating a more modern analysis of the quality of the water based on where it was sourced. The table below provides a brief overview of the aqueducts along with several notes about their use:

Table 5.2.4.*a***:** Aqueduct characteristics and notes about water volume, source, quality and use. Table by Rachel Schleiger (CC-BY-NC) modified from data in Aicher PJ 1995.

Aqueduct	Construction Complete	Volume (units)	Water Source	Water Quality	Notes
Appia	312 BC	31	Springs	Good	All underground except inside walls 70% for civic/imperial uses
Anio Vetus	272-269 BC	74	River	Poor	Used for baths, gardens, and industry
Marcia	144-240 BC	78	Springs	Best	Pure/cold/hard water Supplied baths
Tepula	126-125 BC	7	Streams	Good	Warm water (60F)
Julia	33 BC	20	Springs	Good	N/A
Virgo	22-19 BC	41	Marsh	Good	Almost all underground, some along channel Supplied baths
Alsietina	2 BC	7	Lake	Poor	Build to supply basin for mock sea battles
Claudia	38-52 AD	76	Springs	Good	Built several branches in the city
Anio Novus	38-52 AD	78	River	Okay	Quality was poor until later improved
Traiana	109 AD	47	Springs	Good	N/A
Alexandria	226 AD	9	Springs	Good	Served baths

Questions

1. How many of the 11 aqueducts do you think have drinkable water?
2. Does the above answer surprise you? Why/why not?
3. How many of the aqueducts with drinkable water come from a higher volume aqueduct?
4. Based on what you can observe in the notes, what did the Romans do when water wasn't drinkable?
5. What is the most surprising thing to you about the data presented in the above table? Why?

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 5.2.4: Data Dive- Aqueducts of Rome is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **13.4: Data Dive- Aqueducts of Rome** by Melissa Ha and Rachel Schleiger has no license indicated.

5.2.5: Review

Summary

After completing this chapter you should be able to...

- Describe the different water reservoirs and identify which of these reservoirs are accessible to humans.
- Diagram the water cycle and explain how humans interact with the water cycle.
- Explain how groundwater is acquired from aquifers and how aquifers are recharged.
- Identify the main water-consuming sectors and their relative contributions to water usage.
- Compare physical and economic water scarcity.
- Outline strategies for addressing water shortages.
- Provide examples of water conserving technologies and behaviors.

Water is stored in **water reservoirs**. Although there is much water on Earth, only a fraction of this is freshwater that is accessible to humans. The **water cycle** describes the movement of water through each reservoir and involves **sublimination** and **evapotranspiration**, **condensation**, **precipitation**, **infiltration**, and **surface runoff**. Humans alter the water cycle through land use and diverting water flow. Precipitation—a major control of fresh water availability—is unevenly distributed around the globe. More precipitation falls near the equator, and landmasses there are characterized by a tropical rainforest climate. Less precipitation tends to fall near 30 degrees latitude, where the world's largest deserts are located. **Surface waters**, including lakes are rivers, are an important source of freshwater. Additionally, **groundwater** can be extracted from **aquifers**. Agriculture is the greatest consumer of water, followed by industry and municipal use.

The **water crisis** refers to a global situation where people in many areas lack access to sufficient water or clean water or both. **Physical water scarcity** occurs when water resources are insufficient, but people face **economic water scarcity** when they cannot afford access to water. Partial solutions to the water crisis include dams and reservoirs, rainwater harvesting, aqueducts, desalination, water reuse, and water conservation.

Attribution

Modified by Melissa Ha from Water Availability and Use from *Environmental Biology* by Matthew R. Fisher (licensed under CC BY-NC-SA)

This page titled 5.2.5: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **13.5: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

5.3: Agriculture

Chapter Hook

Steep mountain landscapes aren't an ideal place for farming. Yet, for thousands of years farmers in mountainous regions all over the world have grown crops, even on some of the steepest slopes. To accomplish this effectively, farmers needed to carve gigantic steps across mountainous contours, forming terraces. These terraces create enchanting ripple like patterns across the landscape and make one almost forget that these farms are meant to help feed local and global communities.



Figure 5.3.*a*: Terraced rice fields. Image by PxHere (Public Domain).

Since the transition to agriculture 10,000 years ago, human communities have struggled against the reality that soil suffers nutrient depletion through constant plowing and harvesting (mostly nitrogen loss). The specter of a significant die-off in human population owing to stagnant crop yields was averted in the 1970s by the **Green Revolution**, which, through the engineering of new crop varieties, large-scale irrigation projects, and the massive application of petroleum-based fertilizers to supplement nitrogen, increased staple crop production with such success that the numbers suffering malnutrition actually declined worldwide in the last two decades of the 20th century, from 1.9 to 1.4 billion, even as the world's population increased at 100 times background rates. The prospects for expanding those gains in the new century are nevertheless threatened by the success of industrial agriculture itself. Soil depletion, declining water resources, and the diminishing returns of fertilizer technology—all the products of a half-century of industrial agriculture—have seen increases in crop yields level off.

Thumbnail image - Women farmers planting a rice field in West Sumatra.

Attribution

Modified by Melissa Ha and Rachel Schleiger from The Industrialization of Nature: A Modern History (1500 to the present) from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors. Download for free at CNX. (licensed under CC-BY)

5.3.1: Food Security

5.3.2: Industrial Agriculture

5.3.3: Selective Breeding and Genetic Engineering

5.3.4: Sustainable Agriculture

5.3.5: Data Dive- Landslides and Ag Terraces

5.3.6: Review

This page titled 5.3: Agriculture is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

5.3.1: Food Security

Progress continues in the fight against hunger, yet an unacceptably large number of people lack the food they need for an active and healthy life. In 2019, 690 million people in the world (8.9%) were undernourished (lacked sufficient calories; see The State of Food Security...), and 8.2% of the global population lived in poverty (see Goal 1...). Poverty is internationally defined as living on less than $1.90 per day. Poverty—not food availability—is the major driver of food insecurity. Improvements in agricultural productivity are necessary to increase rural household incomes and access to available food but are insufficient to ensure food security. Evidence indicates that poverty reduction and food security do not necessarily move in tandem. The main problem is lack of economic (social and physical) access to food at national and household levels and inadequate nutrition. Food security not only requires an adequate supply of food but also entails availability, access, and utilization by all—people of all ages, gender, ethnicity, religion, and socioeconomic levels.

From Agriculture to Food Security

Agriculture and food security are inextricably linked. The agricultural sector in each country is dependent on the available natural resources, as well as the politics that govern those resources. **Staple food crops** are the main source of dietary energy in the human diet and include things such as rice, wheat, sweet potatoes, maize (corn), and cassava (figure $5.3.1.a$).



Figure $5.3.1.a$: Cassava tubers are a staple food crop in tropical regions. Image by Renatosjoao (CC-BY-SA).

Food Security

An individual must have access to food in ample quantity and of sufficient nutritional quality at all times to be **food secure**. Those who never have sufficient quality food are **chronically food insecure**. Food security is determined by availability, access, and utilization.

Food availability refers to whether enough food is produced globally to feed the world's population. In fact, enough food is produced globally, but it is not accessible to everyone who needs it. **Food access** refers to the ability to obtain food in ample quantity and of nutritional quality. In the United States, residents of **food deserts** face limited access to nutritious foods, particularly fresh produce (figure $5.3.1.b$). This problem is compounded when residents do not own a vehicle or have access to public transportation to obtain foods from surrounding areas or lack the funds to do so.

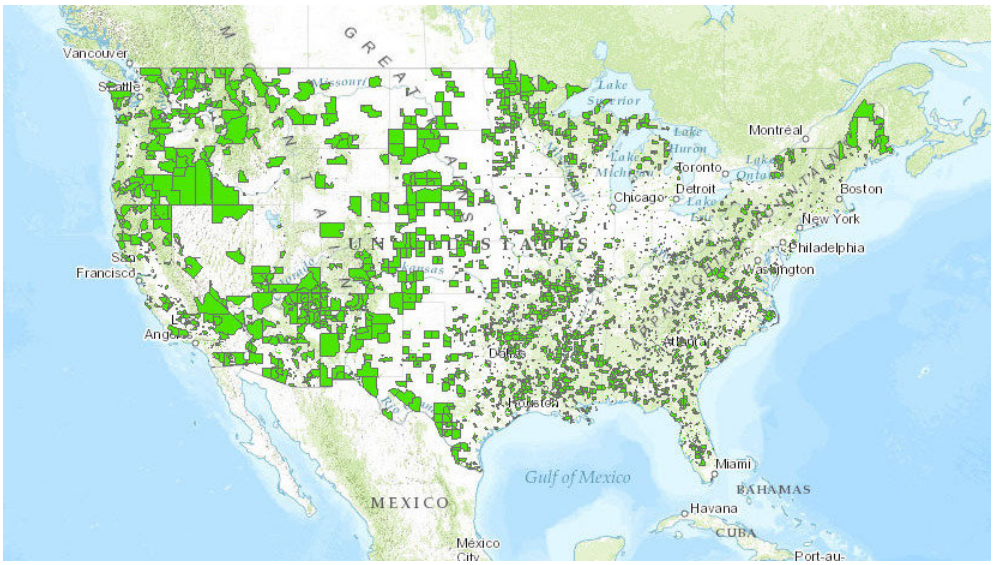


Figure 5.3.1.*b*: Food deserts in the contiguous United States are shaded in bright green. Food deserts are distributed over the United States. Examples of large food deserts are in southeastern Oregon and northeastern Arizona. Image by Economic Research Services, USDA (public domain).

Food utilization essentially translates the food available to a household into nutritional security for its members. One aspect of utilization is analyzed in terms of distribution according to need. Nutritional standards exist for the actual nutritional needs, which differ by gender, age, and life phase (for example, pregnancy), but these "needs" are often socially constructed based on culture. For example, in South Asia evidence shows that women eat after everyone else has eaten and are less likely than men in the same household to consume preferred foods such as meats and fish.

Food insecurity can result in **undernourishment** and **malnutrition** (a lack of essential nutrients; figure 5.3.1.*c*). The economic costs of undernourishment and malnutrition are substantial, potentially costing individuals 10% of their lifetime earnings and costing nations 2-3% of gross domestic product (GDP) in the worst-affected countries (Alderman 2005). This in part due to the reduced ability of under- and malnourished individuals to work. Additionally, without proper nutrition, children can develop disabilities that make it more difficult for them to join to workforce as adults. The economic costs reported here do not even account for the healthcare costs associated with under- and malnutrition.



Figure 5.3.1.*c*: A goiter is an enlarged thyroid gland that can result from deficiency of iodine, an essential nutrient. Iodized salt, commonly used in the United States, is one strategy to prevent iodine deficiencies. Image by Almazi/OpenStax (CC-BY). Access for free at openstax.org.

Obesity

Obesity means having excess body fat resulting in a body mass index (BMI) of 30 or higher, and **overweight** individuals have a BMI of 25 or higher (figure 5.3.1.*d*). BMI is a common metric for obesity because it is determined based on weight and height; however, an individual with a big bone structure and/or high muscle mass could have a high BMI without facing the health risks associated with obesity. Likewise, an individual with excess body fat but little muscle mass might have a "healthy" BMI, but their excess body fat could still compromise their health. For these reasons, some experts rely on other metrics, such as relative fat mass (RFM). Relative fat mass is equal to height divided by waist circumference multiplied by 76 for females or 64 for males. An RFM of 32% or higher for females and 25% or higher for males is considered obese. Harvard Chan School of Public Health provides a thorough discussion of the advantages and disadvantages of various obesity metrics.

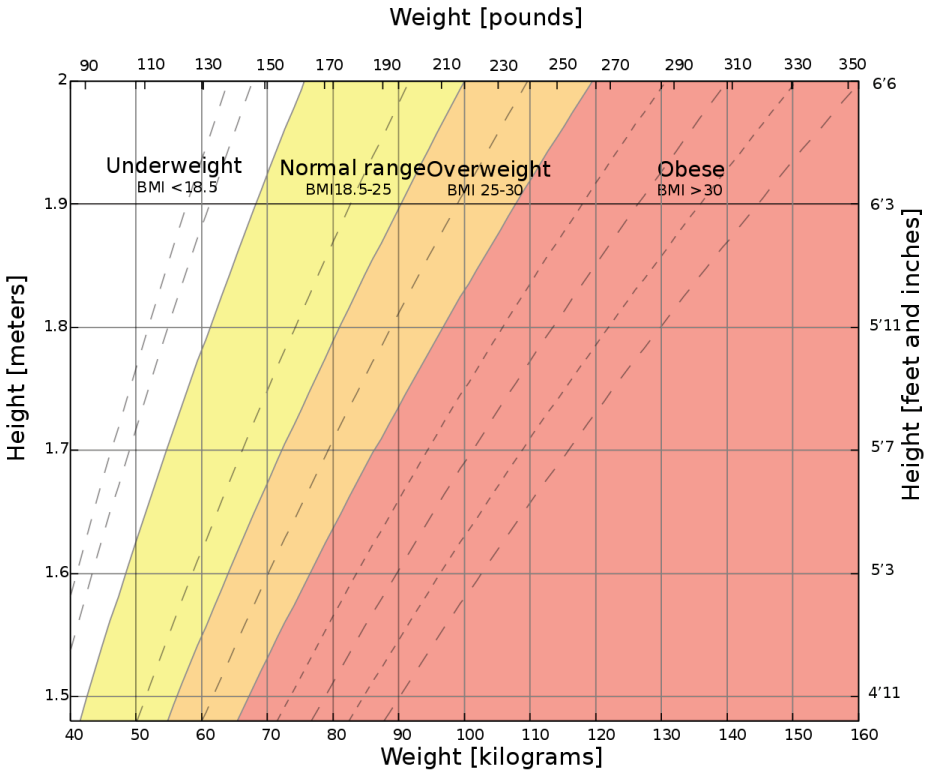


Figure 5.3.1.*d*: This graph shows BMI categories based on height and weight. Height in feet and inches is on the right axis, and weight in pounds is on the top axis. There are four BMI categories: underweight (< 18.5), normal (18.5-25), overweight (25-30), and obese (> 30). BMI is calculated as mass divided by height squared (kg/m^2). Image by InvictaHOG (public domain).

Obesity has become a significant global health challenge, yet is preventable and reversible. Since the 1970s, obesity rates have rapidly increased, resulting in a global obesity epidemic. Thirty-nine percent of the world's adults, or 1.9 billion people, are overweight in 2016. Of these, 650 million are obese (Obesity and Overweight). Obesity was linked to 4.7 million premature deaths in 2017 (about 8% of all deaths that year; Hannah/Our World in Data). The economic cost of obesity has been estimated at $2 trillion (Tremmel et al.).

Initially centered in developed countries, the obesity epidemic now affects countries of all economic statuses. In low-income countries, this results in a triple burden of undernourishment, malnutrition, and obesity. There is significant variation by region; some have very high rates of undernourishment and low rates of obesity, while in other regions the opposite is true (figure 5.3.1.*e*).

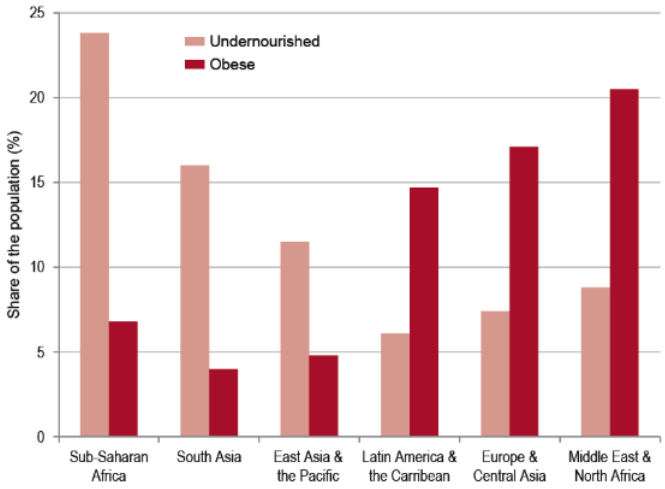


Figure 5.3.1.*e*: Obesity (red bars on the right) and undernourishment (pink bars on the left) by region. The share of the population (%) that is undernourished is approximately 24% in Sub-Saharan Africa while approximately 7% of the population in this region is obese. For South Asia 16% of the population is undernourished, and 4% is obese. These percentages 11% and 5% for East Asia and the Pacific, 5% and 15% for Latin America and the Caribbean, 7% and 17% for Europe and Central Asia, and 9% and 20% for the Middle East and North Africa, respectively.

Obesity accounts for a growing level and share of worldwide noncommunicable diseases, including diabetes, heart disease, and certain cancers that can reduce quality of life and increase public health costs of already under-resourced developing countries. Driven primarily by increasing availability of processed, affordable, and effectively marketed food, the global food system is falling short with rising obesity and related poor health outcomes. Due to established health implications and rapid increase in prevalence, obesity is now a recognized major global health challenge.

Supplemental Reading

The United Nations explains global food supply and strategies for addressing world hunger here using a variety of informational graphics and videos.

References

Alderman, H. 2005. Linkages between Poverty Reduction Strategies and Child Nutrition: An Asian Perspective. *Economic and Political Weekly*, *40*(46), 4837-4842. Accessed 2021-01-03.

Hannah Ritchie. 2017. Obesity. Our World in Data. Accessed 2021-01-03.

Goal 1: End poverty in all its forms everywhere. Sustainable Development. Department of Economic and Social Affairs. United Nations. Accessed 2021-01-03.

Obesity and Overweight. 2020. World Health Organization. Accessed 2021-01-03.

The State of Food Security and Nutrition in the World in 2020. Food and Agriculture Organization of the United Nations. Accessed 2021-01-03.

Tremmel, M., Gerdtham, U. G., Nilsson, P. M., & Saha, S. 2017. Economic Burden of Obesity: A Systematic Literature Review. *International journal of environmental research and public health*, *14*(4), 435. Accessed 2021-01-03.

Attribution

Modified by Melissa Ha from Food Security from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 5.3.1: Food Security is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **14.1: Food Security** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

5.3.2: Industrial Agriculture

Also known as conventional agriculture, **industrial agriculture** is a method of farming that involves use of synthetic fertilizer, synthetic pesticides, and machinery. Industrial agriculture is dependent on large investments in mechanized equipment powered mostly by fossil fuels. In the case of livestock, most production comes from systems where animals are highly concentrated and confined.

Synthetic Pesticides

Pests are organisms that occur where they are not wanted or that cause damage to crops, humans or other animals (figure 5.3.2.*a*). Thus, the term "pest" is a highly subjective term. A **pesticide** is a term for any substance intended for preventing, destroying, repelling, or mitigating any pest. Though often misunderstood to refer only to insecticides, the term pesticide also applies to **herbicides** (weed killers), fungicides, and various other substances used to control pests. Chemical pesticides can be effective, fast acting, adaptable to all crops and situations. When first applied, pesticides can result in impressive production gains of crops. However, despite these initial gains, excessive use of pesticides can be ecologically unsound (see Disadvantages of Industrial Agriculture). By their very nature, most pesticides create some risk of harm—pesticides can cause harm to humans, animals, and/or the environment because they are designed to kill or otherwise adversely affect living things. At the same time, pesticides are useful to society because they can kill potential disease-causing organisms and control insects, weeds, worms, and fungi.



Figure 5.3.2.*a*: The cotton boll weevil is considered a major pest because of the damage it does to cotton plants. Image by ARS USDA (public domain).

Monoculture Farming

Industrial agriculture employs **monoculture** farming, which involves growing only one species of crop in a large area (figure 5.3.2.*b*). The plants in a monoculture are evenly spaced and have the same planting, irrigation, fertilizer, harvesting, etc. requirements, and this uniformity results in efficient use of farming machinery. Often, the crop is a **high yield variety** (high yielding variety; HYV), which is produced through selectively breeding individual plants with desirable characteristics. Compared to traditional crops, HYVs can produce more per unit area of land. Monoculture farming compromises genetic and species diversity in agriculture, however, promoting the spread of pests and risking that all the plants for hundreds of acres could be susceptible to the same disease (see Genetic Diversity).



Figure 5.3.2.*b*: Monocultures like this potato farm (left) and cotton farm (right) make it easier to use farming equipment; however, this reduces genetic diversity and consumes fossil fuels. Left image by NightThree (CC-BY) and right image by Kimberly Vardeman (CC-BY).

Advantages of Industrial Agriculture

Industrial agriculture has delivered tremendous gains in productivity and efficiency. Food production worldwide has overall risen since the 1940s (figure 5.3.2.*c*); the World Bank estimates that between 70 percent and 90 percent of the recent increases in food production are the result of industrial agriculture rather than greater acreage under cultivation (figure 5.3.2.*d*). United States consumers have come to expect abundant and inexpensive food. Additionally, the use of synthetic pesticides ensures that produce is relatively free of blemishes (figure 5.3.2.*e*).

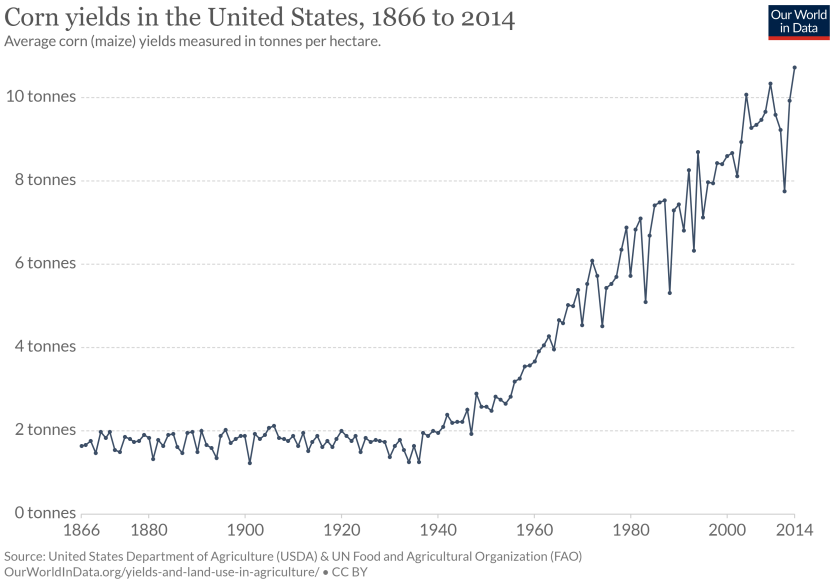


Figure 5.3.2.*c*: Corn yields in metric tons (tonnes; 1000 kilograms) per hectare over time. While yields have fluctuated from year to year, they have overall increased since the 1940s. This improving agricultural efficiency is partly due to industrial agricultural practices. Image by Hannah Ritchie and Max Roser/Our World in Data (CC-BY).

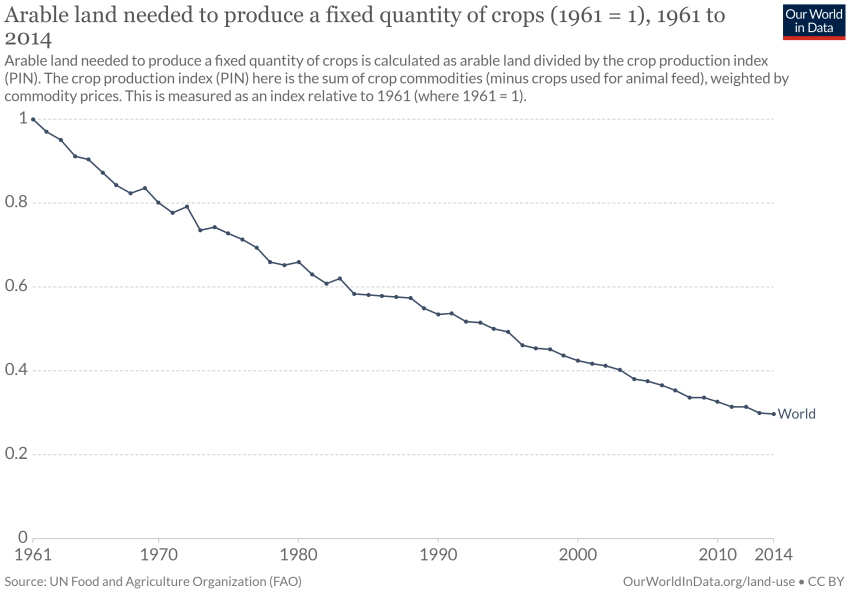


Figure 5.3.2.*d***:** The arable (farmable) land required to produce a fixed amount of a crop has decreased over time due to improved agricultural efficiency. Image by Hannah Ritchie and Max Roser/Our World in Data (CC-BY).



Figure 5.3.2.*e*: Many of us carefully inspect produce at the grocery store and select those without blemishes. This visually appealing produce results from synthetic pesticides use. Image by Amanda Mills (public domain).

Industrial agriculture expands where crops can be cultivated. Applying synthetic fertilizers to poor soils can make them fertile. Additionally, irrigation systems promote high agricultural productivity in regions and during seasons that would otherwise only support the most drought-tolerant species. For example, about 25% of the food consumed in the United States is produced in California's Central Valley, which receives little rain in the summer months (figure 5.3.2.*f*).

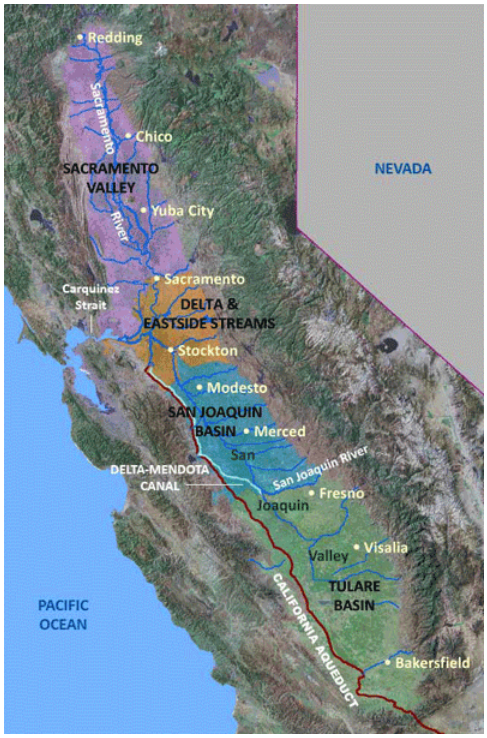


Figure 5.3.2.*f*: The four regions of California's Central Valley from north to south are the Sacramento Valley, Delta and Eastside Streams, San Joaquin Basin, and Tulare Basin. Image by USGS (public domain).

Disadvantages of Industrial Agriculture

Economically, the U.S. agricultural sector includes a history of increasingly large federal expenditures. Also observed is a widening disparity among the income of farmers and the escalating concentration of **agribusiness**—industries involved with manufacture,

processing, and distribution of farm products—into fewer and fewer hands. Market competition is limited and farmers have little control over prices of their goods, and they continue to receive a smaller and smaller portion of consumer dollars spent on agricultural products.

Economic pressures have led to a tremendous loss to the number of farms, particularly small farms, and farmers during the past few decades. More than 155,000 farms were lost from 1987 to 1997. Economically, it is very difficult for potential farmers to enter the business today because of the high cost of doing business. Farming equipment is expensive, and the industrial agricultural system favors large corporate farms, which can more easily invest in this equipment. While this equipment promotes efficiency, it also reduces the number of jobs in agriculture (figure 5.3.2.*g*).

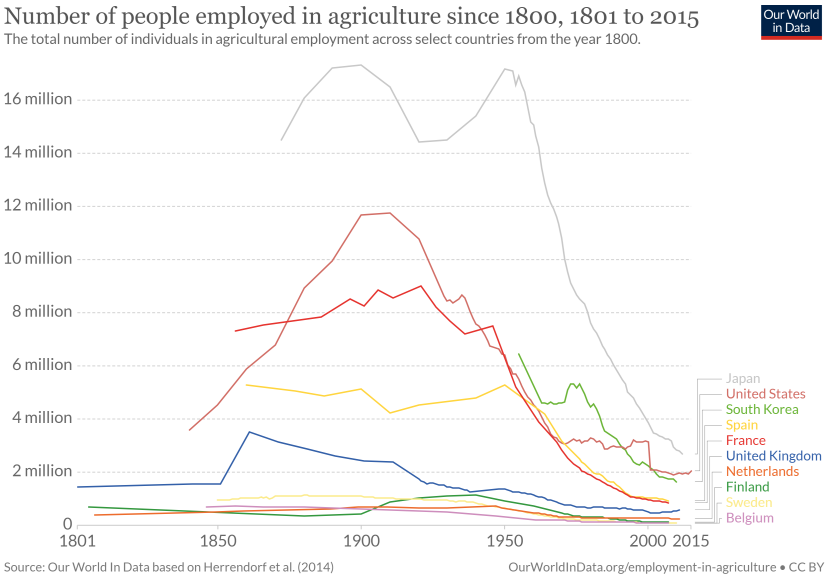


Figure 5.3.2.*g*: The number of people working in agriculture in selected countries over time (from 1801-2015). Generally, the number of agricultural workers has declined as farming practices increased in efficiency and as farming machinery completed tasks that were once manual. This is both an advantage because people are free to pursue other careers and a disadvantage because of the reduction in jobs. In Japan, the number of agricultural workers peaked above 16 million around 1900 and again around 1960. In the United States, it peaked to nearly 12 million around 1900. France followed a similar pattern, peaking around 1925 at approximately 9 million workers. There were over 6 million agricultural workers in South Korea in 1950, and this has overall declined (although there was a spike around 1980). In Spain the number of workers was over 5 million in 1950 and has since decreased. In the United Kingdom, the number of agricultural workers peaked around 1960 at about 3.5 million. The other countries shown (Netherlands, Finland, Sweden, and Belgium) have consistently had under 1 million agricultural workers. Image by Max Roser/Our World in Data (CC-BY).

The mechanical farming equipment required for industrial agriculture relies on fossil fuels. Burning fossil fuels releases air pollution, including greenhouse gases, which cause climate change. Additionally, the synthetic fertilizers and pesticides used in industrial agriculture are produced from fossil fuels. Furthermore, agriculture areas rarely contain as much biomass (organic matter that comprises living things) as biodiverse, intact ecosystems. Biomass is an important carbon sink because its organic molecules store carbon that could otherwise be in the atmosphere as carbon dioxide or methane, both greenhouse gases. Soils with a substantial organic component are also carbon sinks. Industrial agricultural practices that deplete these soils (see below) contribute to climate change, but alternative practices can promote this benefit of healthy soils (see Sustainable Agriculture).

While synthetic fertilizers provide crops with the necessary nutrients for high productivity, excess nutrients from fertilizers can enter bodies of water through runoff. This can lead to a bloom of algae or photosynthetic bacteria in lakes, rivers, and bays in a process called **eutrophication**. These photosynthetic microorganisms sometimes produce toxins that kill aquatic animals and can even harm the humans who consume these animals. Furthermore, they block aquatic plants from accessing light. Eventual decomposition of algal blooms requires oxygen, which results in **hypoxia** (low levels of dissolved oxygen), further harming aquatic species. Additionally, industrial agriculture pollutes water with pesticides and sediments. Pesticides are not only widespread in surface waters, but pesticides from every chemical class have been detected in groundwater.

Industrial agricultural practices often deplete soil quality. For example, wind and water erosion of exposed topsoil removes particles and nutrients from the soil (which also contributes to eutrophication and sediment pollution). Tilling (mixing the soil) and overgrazing of livestock exacerbates erosion. When livestock remove too much vegetation from an area, plant roots no longer anchor the soil in place, and erosion ensues. Farming equipment compacts the soil. This limits infiltration of water into the soil (decreases soil permeability), makes it more difficult for roots to penetrate the soil, and hinders gas exchange in roots. In extreme cases, soil has eroded and lost its ability to retain water that once arable (farmable) land becomes a desert-like (**desertification**; see Soil Degradation; figure $5.3.2.h$).



Figure $5.3.2.h$: Overgrazing depleted the soil in the Rukwa Plains of southwestern Tanzania. Image by Lichinga (CC-BY-SA).

Irrigation can lead to **salinization** (increased salinity). While the water itself eventually evaporates, is transpired, or drains from the soil, the minerals dissolved in the water may remain in the soil. Over time, these minerals can accumulate to levels that are toxic for most plants. As discussed in Water Usage, irrigation for agriculture accounts for 69% of global water usage. Withdrawal of sufficient water for irrigation contributes to water shortages and diverts water from ecosystems (see Aqueducts).

Repeated pesticide application exerts selective pressure on insects, fungi, and other crop pests to evolve pesticide resistance. By chance, some individuals in a pest population may carry gene versions that confer pesticide resistance. When these populations are exposed to pesticides, individuals with these gene versions are the most likely to survive and reproduce. They then pass these genes to their offspring, and pesticide resistance becomes more and more common in the population over time (figure $5.3.2.i$). At this point, pesticides may need to be applied at greater concentrations or more frequently to achieve the same effect. If pesticide that was initially used is no longer effective at all, the farmer would be forced to find a different pesticide. Over 400 insects and mite pests and more than 70 fungal pathogens have become resistant to one or more pesticides.

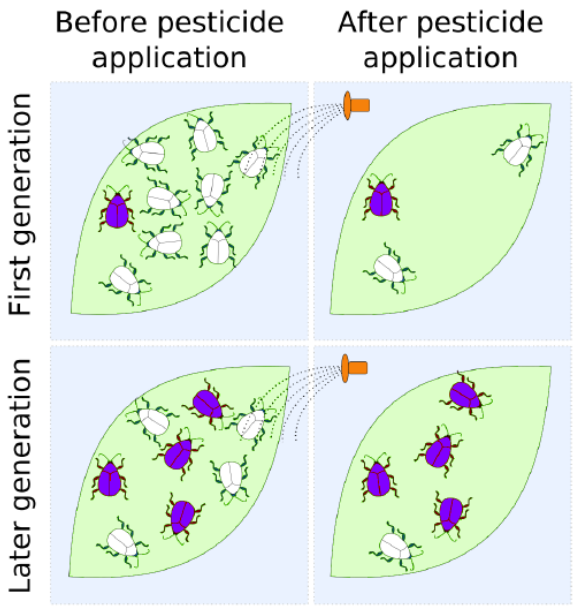


Figure 5.3.2.*i*: Pesticide application can exert selection for the evolution of pesticide resistance. In this hypothetical example, the purple beetle has a gene version that makes it resistant to pesticides (upper left). Pesticides are applied, killing most of the susceptible individuals (white beetles; upper right). In a later generation, purple, resistant beetles are more common because they had a better chance of surviving and reproducing (lower left). Upon another pesticide application, an even greater proportion of the population is resistant. Image modified from Delldot (CC-BY-SA).

The human health and environmental costs of pesticide use have been unevenly distributed. Despite the fact that the lion's share of chemical pesticides are applied in developed countries, 99 percent of all pesticide poisoning cases occur in developing countries where regulatory, health and education systems are weakest. Many farmers in developing countries overuse pesticides and do not take proper safety precautions because they do not understand the risks and fear smaller harvests. Making matters worse, developing countries seldom have strong regulatory systems for dangerous chemicals; pesticides banned or restricted in industrialized countries are used widely in developing countries. Farmers' perceptions of appropriate pesticide use vary by setting and culture. Prolonged exposure to pesticides has been associated with several chronic and acute health effects like non-Hodgkin's lymphoma, leukemia, as well as cardiopulmonary disorders, neurological and hematological symptoms, and skin diseases.

Pesticides have also placed stresses on pollinators and other beneficial insect species. They can lead to the destruction of natural enemies, which would regulate pest populations in an intact ecosystem. For example, once apple growers began controlling pests with the pesticide dichlorodiphenyltrichloroethane (DDT), which has since been banned, they quickly found their orchards being attacked by scale insects and mites. The reason: DDT had killed off their natural enemies. More generally, pollution from industrial agriculture along with habitat loss due to converting healthy ecosystems into agricultural fields is a major threat to biodiversity. As mentioned previously, monoculture farming limits biodiversity in crops.

With respect to animal agriculture, keeping livestock are kept in close quarters increases the risk of disease spread among them. For these reason, food animals are commonly given antibiotics, either preventatively or to treat existing infections. As discussed in the Infectious Diseases section, the overuse of antibiotics (whether in humans or farm animals) increases selection for the evolution of antibiotic-resistant strains of bacteria or other disease-causing organisms. Infections by antibiotic-resistant strains are more difficult to treat.

As with many industrial practices, potential health hazards are often tied to farming practices. Under research and investigation currently is the sub-therapeutic use of antibiotics in animal production, and pesticide and nitrate contamination of water and food. Farm worker health is also a consideration in all farming practices.

Attributions

Modified by Melissa Ha from the following sources:

- Conventional Agriculture and Pests and Pesticides from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

- Insecticides from *Biology* by John W. Kimball (licensed under CC-BY)
- Sustainable Agriculture: Definitions and Terms by Mary V. Gold, Alternative Farming Systems Information Center from National Agricultural Library, USDA (public domain)

This page titled 5.3.2: Industrial Agriculture is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **14.2: Industrial Agriculture** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

5.3.3: Selective Breeding and Genetic Engineering

The development of a new crop variety is an example of **agricultural biotechnology**, a range of tools that include both traditional breeding techniques and more modern lab-based methods. Traditional methods date back thousands of years, whereas biotechnology uses the tools of genetic engineering developed over the last few decades.

Selective Breeding (Artificial Selection)

Nearly all the fruits and vegetables found in your local market would not occur naturally. In fact, they exist only because of human intervention that began thousands of years ago. Humans created the vast majority of crop species by using traditional breeding practices on naturally-occurring, wild plants. These practices rely upon **selective breeding** (artificial selection), human-facilitated reproduction of individuals with desirable traits. For example, high yield varieties were produced through selective breeding. Traditional breeding practices, although low-tech and simple to perform, have the practical outcome of modifying an organism's genetic information, thus producing new traits.

Selective breeding is limited, however, by the life cycle of the plant and the genetic variants that are naturally present. For example, even the fastest flowering corn variety has a generation time of 60 days (the time required for a seed to germinate, produce a mature plant, get pollinated, and ultimately produce more seeds) in perfect conditions. Each generation provides an opportunity to selectively breed individual plants and generate seeds that are slightly closer to the desired outcome (for example, producing bigger, juicier kernels). Furthermore, if no individuals happen to possess gene variants that result in bigger, juicier kernels, it is not possible to artificially select this trait. Finally, traditional breeding shuffles all of the genes between the two individuals being bred, which can number into the tens of thousands (maize, for example, has 32,000 genes). When mixing such a large number of genes, the results can be unpredictable.

An interesting example is maize (corn). Biologists have discovered that maize was developed from a wild plant called teosinte. Through traditional breeding practices, humans living thousands of years ago in what is now Southern Mexico began selecting for desirable traits until they were able to transform the plant into what is now known as maize (figure 5.3.3.a). In doing so, they permanently (and unknowingly) altered its genetic instructions.

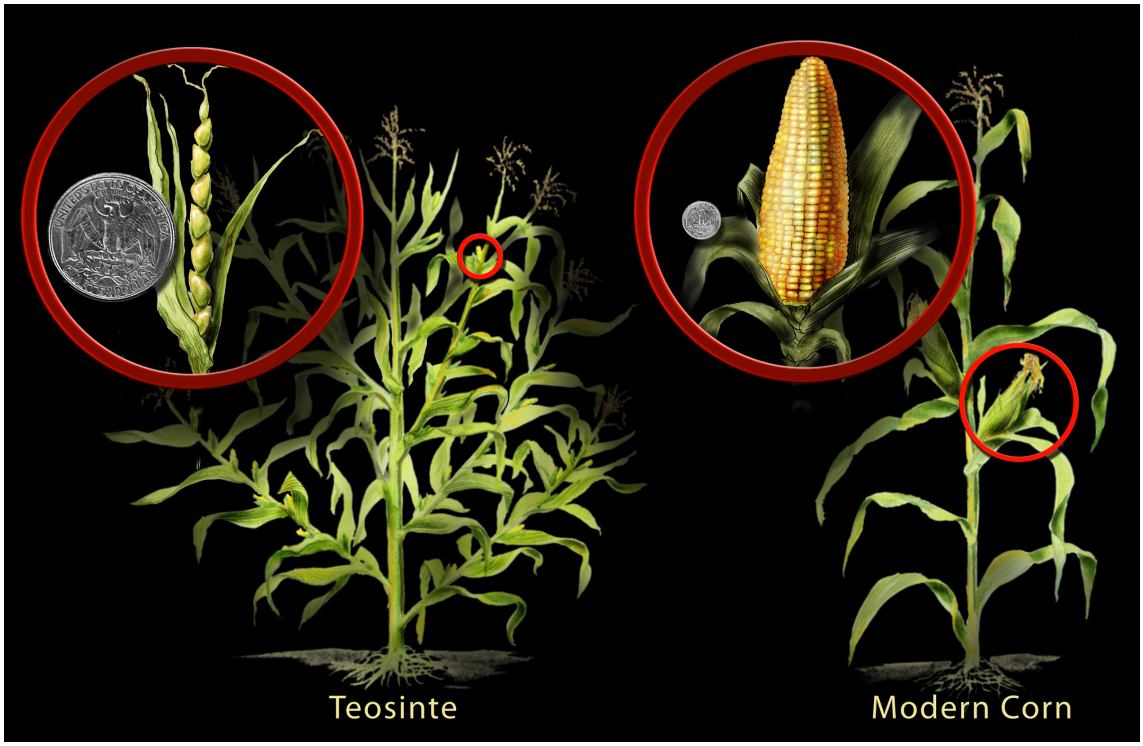


Figure 5.3.3.a: A wild grass called teosinte was genetically modified through selective breeding to produce what is now known as maize (corn). This process of transformation started thousands of years ago by indigenous people of what is now Mexico. Image by Nicolle Rager Fuller/National Science Foundation (public domain).

This history of genetic modification is common to nearly all crop species. For example, cabbage, broccoli, Brussel sprouts, cauliflower, and kale were all developed from a single species of wild mustard plant (figure 5.3.3.*b*). Wild nightshade was the source of tomatoes, eggplant, tobacco, and potatoes, the latter developed by humans 7,000 – 10,000 years ago in South America.



Figure 5.3.3.*b*: *Brassica oleracea* is a plant in the mustard family and is known as wild cabbage. From it were developed many familiar crops, such as cauliflower, broccoli, Brussel sprouts, and of course, cabbage. Image by Kulac (CC-BY-SA).

Genetic Engineering

Genetic engineering is the process of directly altering an organism's DNA to produce the desired crops more rapidly than selective breeding. Because genes can be obtained from other species or even synthesized in the lab, scientists are not limited by existing genetic variation within a crop species (or closely related species with which they can be crossed). This broadens the possible traits that can be added to crops. Modern genetic engineering is more precise than selective breeding in the sense that biologists can modify just a single gene. Also, genetic engineering can introduce a gene between two distantly-related species, such as inserting a bacterial gene into a plant (figure 5.3.3.*c*).

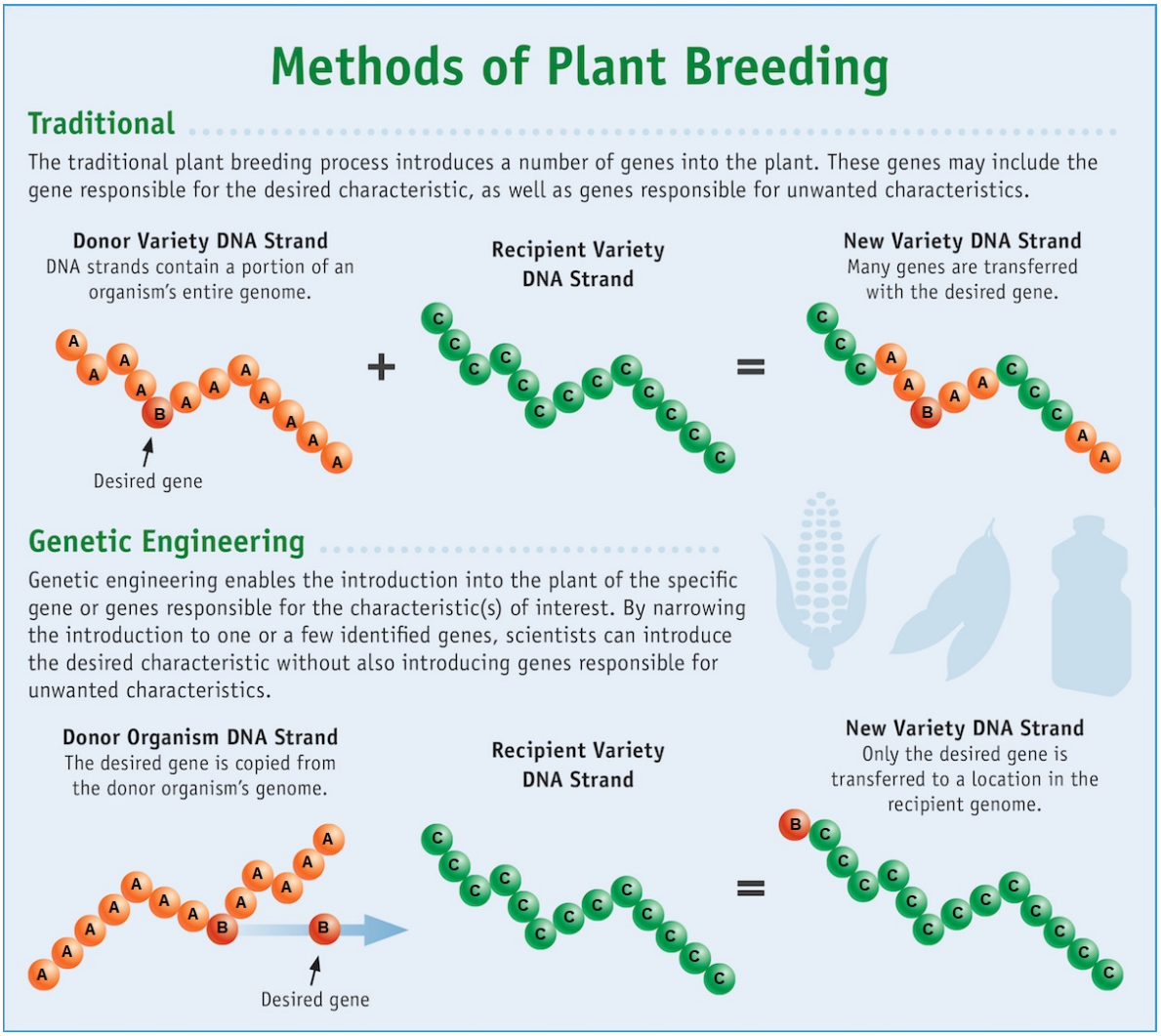


Figure 5.3.3.*c*: Both selective (traditional) breeding and modern genetic engineering produce genetic modifications. Genetic engineering allows for fewer and more precise genetic modifications. The traditional plant breeding process introduces a number of genes into the plant. These genes may include the gene responsible for the desired characteristic, as well as genes responsible for unwanted characteristics. In traditional breeding, the donor variety DNA strand (A) and recipient variety DNA strand (C) are combined to produce the new variety DNA strand. The donor DNA strand contains a portion of an organism's entire genome, including the desired gene (B). In the new variety DNA strand, many genes are transferred with the desired gene. Genetic engineering enables the introduction into the plant of the specific gene or genes responsible for the characteristics of interest. By narrowing the introduction to one of a few identified genes, scientists can introduce the desired characteristic without also introducing genes responsible for unwanted characteristics. In genetic engineering, the desired gene (B) is copied from the donor organism's genome (A). The desired gene from the donor organism DNA strand and the recipient variety DNA strand (C) are combined to produce the new variety DNA strand. Only the desired gene is transferred to a location in the recipient genome. Modified from Michael J. Ermarth/FDA (public domain).

Genetically modified organisms (**GMOs**) are those that have had their DNA altered through genetic engineering. Genetically modified crops are sometimes called genetically engineered (GE) crops. **Transgenic organisms** are a type of genetically modified organism that contains genes from a different species. Because they contain unique combinations of genes and are not restricted to the laboratory, transgenic plants and other GMOs are closely monitored by government agencies to ensure that they are fit for human consumption and do not endanger other plant and animal life. Because these foreign genes (transgenes) can spread to other species in the environment, particularly in the pollen and seeds of plants, extensive testing is required to ensure ecological stability.

How to Genetically Modify Plant Cells

DNA can be inserted into plant cells through various techniques. For example, a **gene gun** propels DNA bound to gold particles into plant cells. (DNA is negatively charge and clings to positively charged gold.) A more traditional approach employs the plant pathogen ***Agrobacterium tumefaciens*** (figure 5.3.3.*d*). Ordinarily, this bacterium causes crown gall disease in plants by inserting

a circular piece of DNA, called the **Ti plasmid**, into plant cells. This DNA incorporates into plant chromosomes, giving them genes to produce the gall (figure 5.3.3.*e*), which provides a home for the bacterial pathogen.

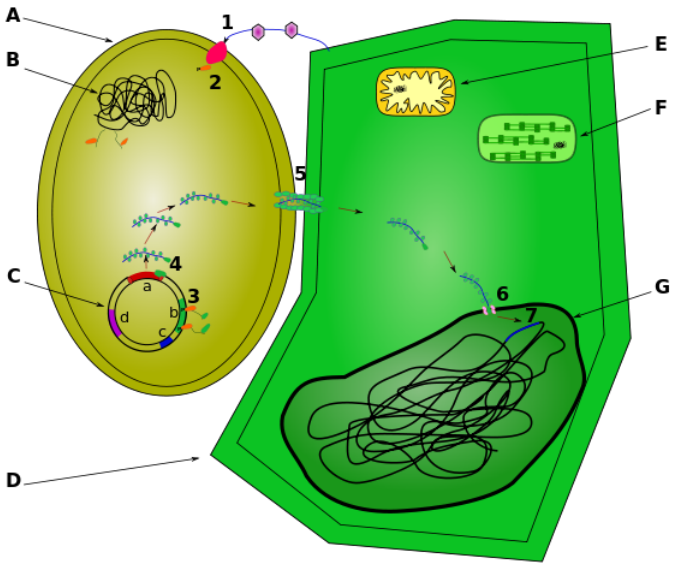


Figure 5.3.3.*d*: An *Agrobacterium tumefaciens* cell (A) inserts the Ti plasmid (C), which contains T-DNA (a) into a plant cell (D). The T-DNA eventually enters the plant nucleus (G), where the plant DNA is stored. Image by Chandres (CC-BY-SA).



Figure 5.3.3.*e*: A crown gall caused by *Agrobacterium tumefaciens*. This bacterium inserts DNA into plant cells, which causes them to produce the crown gall. This process can be adjusted by scientists to insert genes that code for desirable traits into crop species. Image by Scott Nelson (public domain).

Scientists alters the process by which *Agrobacterium* infects and genetically alter plant cells to produce genetically modified plants with agriculturally beneficial traits as follows (figure 5.3.3.*f*):

1. T-DNA, which codes for the crown gall is removed from the Ti plasmid, and genes for desired traits are added.
2. The modified plasmid is then added back to *Agrobacterium.*
3. *Agrobacterium* infects undifferentiated plant cells (stem cells that can develop into any part of the plant; figure 5.3.3.*g*).

4. The modified plant cells are given hormones to produce the entire plant.

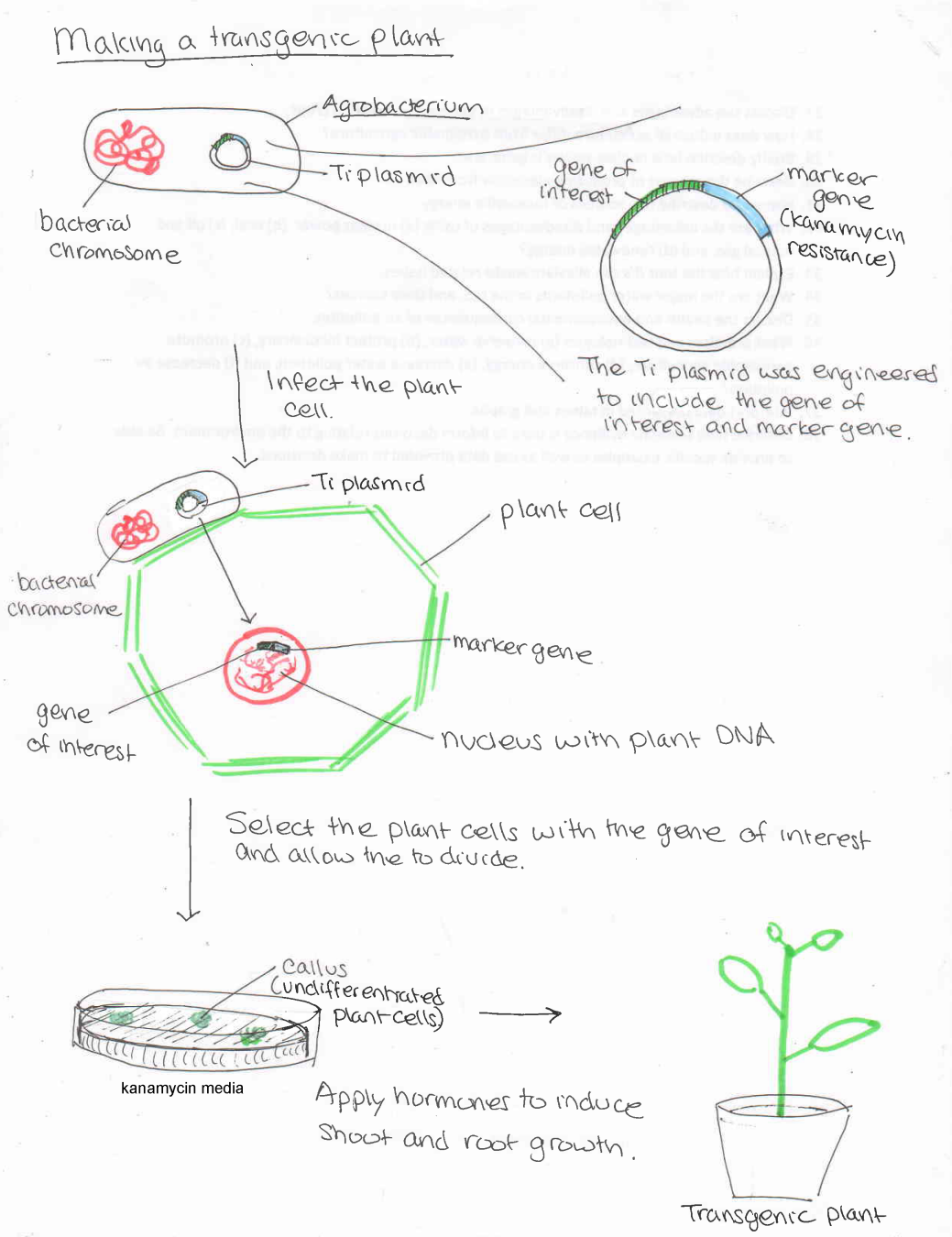


Figure 5.3.3.f: The process of producing a transgenic plant using *Agrobacterium*. The Ti plasmid was engineered to include the gene of interest and marker gene, which will later indicate which cells are transgenic. A rectangle represents *Agrobacterium*. It contains regular bacterial DNA and the Ti plasmid, represented by a circle. The T-DNA (which provides instructors for producing a crown gall) is removed, and a transgene (gene of interest) and maker gene are inserted instead. The gene of interest is colored green, and the marker gene (for resistance to the antibiotic kanamycin) is colored blue. Next, plant cells are infected with *Agrobacterium*. The bacterial cell inserts the Ti plasmid into a polyhedral plant cell. The transgene is transferred to the plant cell nucleus, which contains regular plant DNA. Next, the plant cells with the gene of interest are selected and allowed them to divide. This is done by culturing the plant cells on kanamycin media in a petri dish. Calluses, clusters of undifferentiated plant cells, are produced. Finally, hormones are applied to induce shoot and root growth. An arrow points from the petri dish to a potted transgenic plant. Image by Melissa Ha (CC-BY).

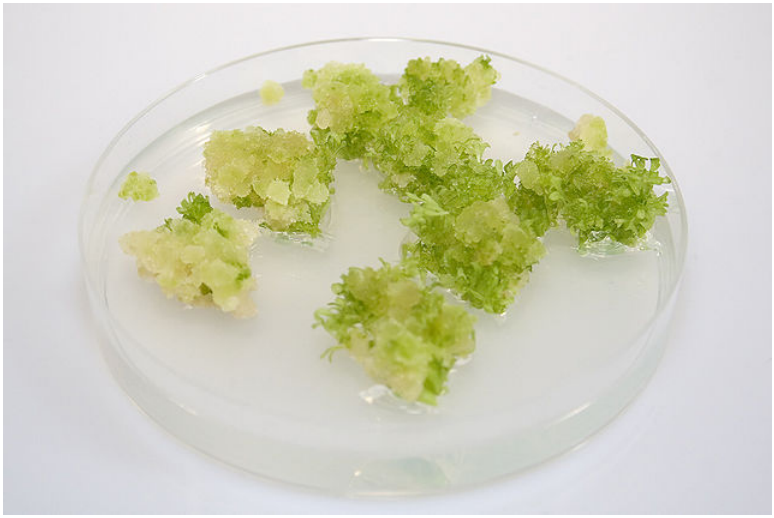


Figure 5.3.3.*g*: Individual plant cells are first genetically engineered with *Agrobacterium*. They then grow into **calluses** (blobs of undifferentiated plant cells) and are given hormones to induce root and shoot development. Because all the cells in the plant that eventually grows descended from a single genetically engineered cell, the entire plant is transgenic. Image by Igge (CC-BY-SA).

Examples of Genetically Modified Crops

Many genetically modified crops have been approved in the U.S. and produce our foods. The first genetically modified organism approved by the U.S. Food and Drug Administration (FDA) in 1994 was **Flavr Savr™ tomatoes**, which have a longer shelf life (delayed rotting) because a gene responsible for breaking down cells in inhibited. Flavr Savr tomatoes are genetically modified (because their DNA has been altered) but not trasgenic (because they do not contain genes from another species). The Flavr Savr tomato did not successfully stay in the market because of problems maintaining and shipping the crop. **Golden rice** produces β-carotene, a precursor to vitamin A (figure 5.3.3.*h*; β-carotene is also in high concentrations in carrots, sweet potatoes, and cantaloupe, giving them their orange color.) **Roundup Ready®** corn, cotton, and soybeans are resistant to this common herbicide, making it easier to uniformly spray it in a field to kill the weeds without harming the crops (figure 5.3.3.*i*).



Figure 5.3.3.*h*: Golden rice has an orange-yellow color because it contains up to 35 µg β-carotene (a precursor to vitamin A) per gram of rice, which could prevent millions of cases of blindness worldwide. Image by International Rice Research Institute (CC-BY).



Figure 5.3.3.*i*: Roundup is a common herbicide. Roundup Ready® plants are genetically modified to resist roundup, meaning the herbicide does not kill them. This allows farmers to uniformly spray Roundup, killing weeds without harming the crops. Image by Mike Mozart (CC-BY).

Crops have also been engineered to produce insecticides. *Bacillus thuringiensis* (Bt) is a bacterium that produces protein crystals that are toxic to many insect species that feed on plants. Insects that have eaten Bt toxin stop feeding on the plants within a few hours. After the toxin is activated in the intestines of the insects, death occurs within a couple of days. The gene to produce **Bt toxin** has been added to many crops including corn (figure 5.3.3.*j*), potatoes, and cotton, providing plants with defense against insects.

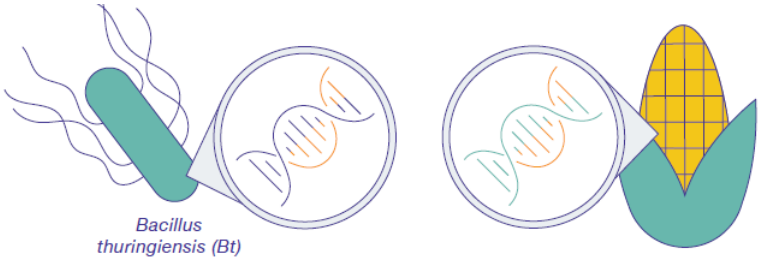


Figure 5.3.3.*j*: Genetic engineering to produce Bt corn. The gene from *Bacillus thuringiensis* that produces Bt toxin is removed and inserted into the corn plant. The Bt corn now produces the insecticide (Bt toxin) and can defend itself against pests. Image by FDA (public domain).

Genetically modified foods are widespread in the United States. For example, 94% of soy crops were genetically modified for herbicide resistance in 2020. Likewise, 8% of cotton and 10% of corn crops were modified for herbicide resistance in addition to the 83% of cotton and 79% of corn crops that were genetically modified in multiple ways.

Genetically modified animals have recently entered the market as well. **AquaAdvantage® salmon** are modified to grow more rapidly and were approved in November of 2015. However, as of March 2021, they have still not been sold due to legal challenges. In 2020, the FDA approved **GalSafe™ pigs** for medicine and food production. These pigs lack a molecule on the outside of their cells that cause allergies in some people.

Advantages of Genetically Modified Crops

Advances in biotechnology may provide consumers with foods that are nutritionally-enriched, longer-lasting, or that contain lower levels of certain naturally occurring toxins present in some food plants. For example, researchers are using biotechnology to try to reduce saturated fats in cooking oils and reduce allergens in foods. Whether these benefits will reach the people who need them most remains to be seen. While cultivating golden rice could address vitamin A deficiency in millions of people, it has not

historically been accessible to these people because it is patented and expensive. Similarly, genetically modified seeds could increase the income of impoverished farmers if they were available at low or no cost, but this is not always the case.

Rainbow and SunUp papayas are a success story of how genetically modified crops can benefit small farmers and the economy in general. In the early 1990s, an emerging disease was destroying Hawaii's production of papaya and threatening to decimate the $11-million industry (figure 5.3.3.*k*). Fortunately, a man named Dennis Gonsalves (figure 5.3.3.*l*), who was raised on a sugar plantation and then became a plant physiologist at Cornell University, would develop papaya plants genetically engineered to resist the deadly virus. By the end of the decade, the Hawaiian papaya industry and the livelihoods of many farmers were saved thanks to the free distribution of Dr. Gonsalves's seeds.

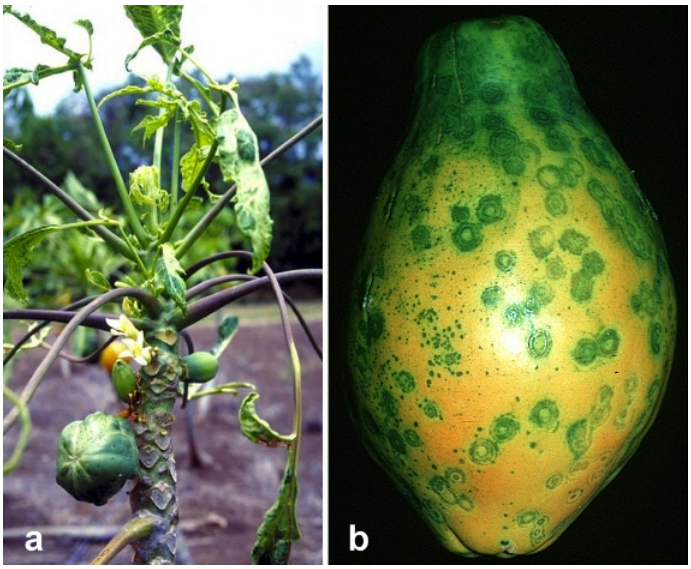


Figure 5.3.3.*k*: The symptoms of papaya ringspot virus are shown on the tree (a) and fruit (b). Image by APS (public domain).



Figure 5.3.3.*l*: Dennis Gonsalves genetically engineered papayas to resist the ringspot virus. Image by ARS USDA (public domain).

The effect of genetically modified crops on the environment depends on the specific genetic modification and which agricultural practices it promotes. For example, Bt crops produce their own insecticides such that external application of these chemicals is unnecessary, reducing the negative impacts of industrial agriculture. Ongoing research is exploring whether crops can be engineered to fix nitrogen in the atmosphere (as some bacteria do) rather than relying on ammonium, nitrites, and nitrates in the soil. If these crops were successfully engineered, they could reduce synthetic fertilizer application and minimize nutrient runoff that leads to eutrophication.

Genetically modified crops may have the potential to conserve natural resources, enable animals to more effectively use nutrients present in feed, and help meet the increasing world food and land demands. In practice, however, countries that use genetically modified crops compared to those that do not only enjoy a slight (or nonexistent) increase in yield.

Disadvantages of Genetically Modified Crops

Social Concerns

Intellectual property rights are one of the important factors in the current debate on genetically modified crops. Genetically modified crops can be patented by agribusinesses, which can lead to them controlling and potentially exploiting agricultural markets. Some accuse companies, such as Monsanto, of allegedly controlling seed production and pricing, much to the detriment of farmers (figure 5.3.3.m).



Figure 5.3.3.m: Protesters in Washington D.C. oppose genetically modified organisms and Monsanto specifically. Image cropped from Sarah Stierch (CC-BY).

Environmental Concerns

Genetically modified crops present several environmental concerns. Monoculture farming already reduces biodiversity, and cultivating genetically modified crops, for which individual plants are quite similar genetically, exacerbates this. The use of Roundup Ready® crops naturally encourages widespread herbicide use, which could unintentionally kill nearby native plants. This practice would also increase herbicide residues on produce. While Bt crops are beneficial in the sense that they do not require external insecticide application, but Bt toxin is spread in their pollen. An early study found that Bt corn pollen may be harmful to monarch caterpillars (figure 5.3.3.n), but only at concentrations that are seldom reached in nature. Follow-up studies found that most of Bt corn grown did not harm monarchs; however, the one strain of Bt corn did was consequently removed from the market.



Figure 5.3.3.n: An early study suggested that pollen containing Bt toxin can harm beneficial and native insects, like this monarch caterpillar. However, without Bt crops, farmers are more likely to spray insecticides, circulating more harmful chemicals than pollen from Bt crops does. Image by Judy Gallagher (CC-BY).

Through interbreeding, or hybridization, genetically modified crops might share their transgenes with wild relatives. This could affect the genetics of those wild relatives and have unforeseen consequences on their populations and could even have implications for the larger ecosystem. For example, if a gene engineered to confer herbicide resistance were to pass from a genetically modified crop to a wild relative, it might transform the wild species into a **superweed** – a species that could not be controlled by herbicide. Its rampant growth could then displace other wild species and the wildlife that depends on it, thus inflecting ecological harm.

Not only could escaped genes alter weedy species, but they could also enter populations of native species. This could make some native species better competitors than they were previously, disrupting ecosystem dynamics. (They could potentially outcompete other native species with which they would otherwise coexist.)

While there is evidence of genetic transfer between genetically modified crops and wild relatives, there is not yet evidence of ecological harm from that transfer. Clearly, continued monitoring, especially for newly-developed crops, is warranted.

The escape of genetically modified animals has potential to disrupt ecosystems as well. For example, if AquaAdvantage salmon were to escape into natural ecosystem, as farmed fish often do, they could outcompete native salmon, including endangered species. Their genetic modification, which facilitates rapid growth, could result in a competitive advantage.

Health Concerns

In addition to environmental risks, some people are concerned about potential health risks of genetically modified crops because they feel that genetic modification alters the intrinsic properties, or essence, of an organism. As discussed above, however, it is known that both traditional breeding practices and modern genetic engineering produce permanent genetic changes. Furthermore, selective breeding actually has a larger and more unpredictable impact on a species's genetics because of its comparably crude nature.

To address these concerns (and others), the US National Academies of Sciences, Engineering, and Medicine (NASEM) published a comprehensive, 500-page report in 2016 that summarized the current scientific knowledge regarding genetically modified crops. The report, titled *Genetically Engineered Crops: Experiences and Prospects*, reviewed more than 900 research articles, in addition to public comments and expert testimony. NASEM's GE Crop Report found "no substantiated evidence of a difference in risks to human health between current commercially available genetically engineered (GE) crops and conventionally bred crops, nor did it find conclusive cause-and-effect evidence of environmental problems from the GE crops." Additionally, the UN's Food and Agriculture Organization has concluded that risks to human and animal health from the use of GMOs are negligible. The scientific consensus on genetically modified crops is quite clear: they are safe for human consumption.

The potential of genetically modified crops to be allergenic is one of the potential adverse health effects, and it should continue to be studied, especially because some scientific evidence indicates that animals fed genetically modified crops have been harmed. NASEM's GE Crop Report concluded that when developing new crops, it is the product that should be studied for potential health and environmental risks, not the process that achieved that product. What this means is, because both traditional breeding practices and modern genetic engineering produce new traits through genetic modification, they both present potential risks. Thus, for the safety of the environment and human health, both should be adequately studied.

Are Genetically Modified Crops the Solution We Need?

Significant resources, both financial and intellectual, have been allocated to answering the question: are genetically modified crops safe for human consumption? After many hundreds of scientific studies, the answer is yes. But a significant question still remains: are they necessary? Certainly, such as in instances like Hawaii's papaya, which were threatened with eradication due to an aggressive disease, genetic engineering was a quick and effective solution that would have been extremely difficult, if not impossible, to solve using traditional breeding practices.

However, in many cases, the early promises of genetically engineered crops – that they would improve nutritional quality of foods, confer disease resistance, and provide unparalleled advances in crop yields – have largely failed to come to fruition. NASEM's GE Crop Report states that while genetically modified crops have resulted in the reduction of agricultural loss from pests, reduced pesticide use, and reduced rates of injury from insecticides for farm workers, they have not increased the rate at which crop yields are advancing when compared to non-GE crops. Additionally, while there are some notable exceptions like golden rice or virus-resistant papayas, very few genetically engineered crops have been produced to increase nutritional capacity or to prevent plant disease that can devastate a farmer's income and reduce food security. The vast majority of genetically modified crops are

developed for only two purposes: to introduce herbicide resistance or pest resistance. Genetically modified crops are concentrated in developed countries, and their availability in developing countries, where they are perhaps most needed, is limited (figure 5.3.3.*o*).

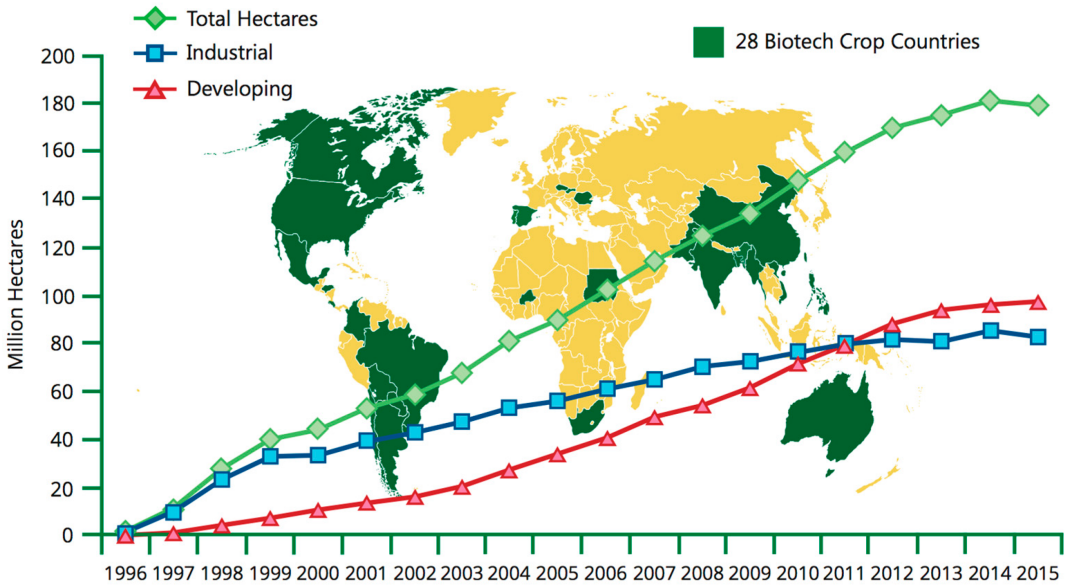


Figure 5.3.3.*o*: Global area of genetically modified crops in millions of hectares 1996–2015. Overall, the hectares of genetically modified crops has increased, and in 2011, developing countries surpassed industrial countries. By 2015, out of the 28 countries that grew biotech crops, 20 were developing, and only eight were developed countries. Latin American, African, and Asian farmers together grew 97.1 million hectares (54%) of the global 179.7 million biotech hectares, whereas industrial countries only planted 83 million hectares or 46%. Image and caption (modified) from Taheri, F., Azadi, H., & D'Haese, M. (2017). A World without Hunger: Organic or GM Crops? *Sustainability*, *9*(4), 580. doi:10.3390/su9040580. (CC-BY)

Suggested Supplementary Reading

NASEM. 2016. Genetically Engineered Crops: Experiences and Prospects.

Attributions

Modified by Melissa Ha from the following sources:

- Biotechnology and Genetic Engineering from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Biotechnology in Medicine and Agriculture from *Concepts in Biology* by OpenStax (licensed under CC-BY)

This page titled 5.3.3: Selective Breeding and Genetic Engineering is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **14.3: Selective Breeding and Genetic Engineering** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

5.3.4: Sustainable Agriculture

Sustainable agriculture is a method of farming that does not deplete natural resources or degrade the environment and can therefore be continued indefinitely. Sustainable farms often rely on components of local ecosystems. For example, they might promote conditions for natural decomposition of wastes or employ natural enemies to control pests through predation or competition. More specifically, the 1977 and 1990 “Farm Bills” describe sustainable agriculture as "an integrated system of plant and animal production practices having a site-specific application that will, over the long term:

- satisfy human food and fiber needs;
- enhance environmental quality and the natural resource base upon which the agricultural economy depends;
- make the most efficient use of nonrenewable resources and on-farm resources and integrate, where appropriate, natural biological cycles and controls;
- sustain the economic viability of farm operations;
- enhance the quality of life for farmers and society as a whole."

Promoting biodiversity is key to sustainable agriculture, and high biodiversity results intact ecosystem services such as nutrient cycling and regulation of pest populations. This is consistent with the goal of sustainable agriculture: to mimic the processes found in natural ecosystems. In contrast to the monocultures of industrial agriculture, **polyculture** farming (growing different species together) is a common practice in sustainable agriculture. This assists with regulating pests and maintaining soil fertility (see below). **Seed banks**, locations at which many different types of seeds are stored, are key to conserving the genetic diversity of crops. These storage sites are cold enough to keep seeds frozen naturally (figure 5.3.4.*a*).



Figure 5.3.4.*a*: The Svalbard Global Seed Vault in Norway. Image by Miksu (CC-BY-SA).

Integrated Pest Management

Integrated Pest Management (IPM) refers to a mix of farmer-driven, ecologically-based pest control practices that seek to reduce reliance on synthetic chemical pesticides. It uses several methods simultaneously to control pest populations. The steps of integrated pest management are to (1) identify true pests, (2) set thresholds and monitor (figure 5.3.4.*b*), and (3) develop an action plan. As many insects, microbes, and other organisms found in an agricultural area have a neutral or beneficial effect, it is not necessary to remove them. True pests are those that are causing economic harm, and they are managed (kept below economically damaging levels) rather than eradicated.



Figure 5.3.4.*b*: A retired Navy commander and landowner inspect an insect trap as part of an Integrated Pest Management (IPM) conservation practice. Image by NRCS (public domain).

An IPM Action Plan draws on four types of control: cultural, mechanical, biological, and chemical. These methods are listed in order of least to most environmental impact and are thus applied in this order. For example, cultural control is attempted first. If that is not effective, mechanical control is added to the plan, and so on. Chemical control is used as a last resort and is deemphasized in an IPM plan. When pesticides must be used, they are selected and applied in a way that minimizes adverse effects on beneficial organisms, humans, and the environment. It is commonly understood that applying an IPM approach does not necessarily mean eliminating pesticide use, although this is often the case because pesticides are often over-used for a variety of reasons.

An Alternative to Spraying: Bollworm Control in Shandong

Farmers in Shandong (China) have been using innovative methods to control bollworm infestation in cotton when this insect became resistant to most pesticides. Among the control measures implemented were:

1. The use of pest resistant cultivars and interplanting of cotton with wheat or maize.
2. Use of lamps and poplar twigs to trap and kill adults to lessen the number of adults.
3. If pesticides were used, they were applied on parts of the cotton plant's stem rather than by spraying the whole field (to protect natural enemies of the bollworm).

These and some additional biological control tools have been effective in controlling insect populations and insect resistance, protecting surroundings and lowering costs.

Cultural Control

Cultural control refers to minimizing the conditions that allow pests to thrive and spread. Examples include alternating which crops are planted each year (**crop rotation**), planting multiple types of crops near each other (**intercropping**; figure 5.3.4.*c*), selecting pest-resistant varieties, and planting pest-free rootstock (underground plant parts). Crop rotation prevents pests that are specialized to a particular type of crop from continuing year after year because their host plants are only available in certain years. Similarly, intercropping spatially limits the spread of pests. **Strip cropping** is a type of intercropping that involves growing different types of crops in alternating rows. Pests may infest one (or several) rows of their host species, but they would have to move past multiple rows of non-hosts to access additional host plants. Additionally, cultural control can involve optimizing irrigation and fertilizer application to promote plant defense and limit disease spread. Cultural control methods can be very effective and cost-efficient and present little to no risk to people or the environment.



Figure 5.3.4.*c*: *Momordica charantia* (bitter melon) intercropped with rice in the Philippines (cultural control). Bitter melon plants emerge from holes in plastic, which is used for mechanical control of weeds. Image by Judgefloro (CC-BY-SA).

Mechanical Control

Mechanical control refers to physically removing pests or excluding them with barriers (figure 5.3.4.*d*). Mechanical controls that remove pests include sticky insect traps, mole traps, and removing weeds by hand. Netting to exclude birds, deer fencing, and weed cloth, plastic weed barriers (figure 5.3.4.*c*) or mulch are additional examples of mechanical control.



Figure 5.3.4.*d*: This solar insect trap is an example of mechanical control. It emits ultraviolet light to attract insects and is recharged using small solar panels. Image by MGK Solar Trap (CC-BY-SA).

Biological Control

Biological control is the use of organisms to reduce pest populations (also see Invasive Species). One biological control strategy involves releasing natural enemies, such as predators, parasites, or parasitoids of the pest organisms (parasitoids are similar to parasites, but they consistently kill their hosts). Home gardeners can also rely on natural enemies by purchasing preying mantises, ladybugs (ladybird beetles), or lacewings for release (figure 5.3.4.*e*). Successful examples include ladybird beetles to depress aphid populations, parasitoid wasps to control whiteflies, and fungi, such as *Trichoderma*, to suppress fungal-caused plant diseases.



Figure 5.3.4.*e*: Lacewings are natural predators of garden and crop pests. Image by artsehn/Pixabay (Pixabay license).

If not used carefully, chemical control can have unintended negative consequences for biological control. In 1887, the cottony cushion scale (native to Australia) was devastating the citrus groves of California. A U.S. entomologist went to Australia to find a natural enemy and came back with the **vedalia beetle**, a species of ladybird beetle. Released in California, the beetle quickly brought the scale under control, at least until 1946. In that year, the pest made a dramatic comeback. This coincided with the first use of the now-banned pesticide dichlorodiphenyltrichloroethane (DDT) in the groves. Not only did DDT kill the target pest insects, but it killed the vedalia beetle as well. Only by altering spray procedures and reintroducing the beetle was the scale insect again controlled.

Another strategy for biological control involves releasing sterile males, which compete with fertile males for mates, ultimately decreasing pest population size. This technique was first applied against the screwworm fly, a serious pest of cattle (figure 5.3.4.*f*). The female flies lay their eggs in sores or other open wounds on the animals. After hatching, the larvae eat the tissues of their host. As they do so, they expose a still larger area to egg laying, often finally killing the host.



Figure 5.3.4.*f*: Thought to be eradicated in the United States by 1966, the screwworm fly, a cattle pest, reemerged in deer at Key Deer National Refuge in Florida in 2016. Sterile males were again released, and the pest was eradicated from Florida in 2017. Image by Judy Gallagher (CC-BY).

Prior to its eradication from the southeastern United States, the screwworm was causing huge annual livestock losses. The sterile male technique involves releasing factory-reared and sterilized flies into the natural population. Sterilization is done by exposing the factory flies to just enough gamma radiation to make them sterile but not enough to reduce their general vigor.

Starting in early 1958, up to 50 million sterilized flies were released each week from aircraft flying over Florida and parts of the adjoining states. Each time a fertile female in the natural population mated with a sterile male, the female layed sterile eggs. Since the females mate only once, her reproductive career was at an end. By early 1959, the pest was totally eliminated east of the Mississippi River. The problem in southwestern states was more challenging because the fly winters in Mexico and could move across the border with each new season. Even so, by expanding the program to include Mexico as well, the screwworm fly was finally eliminated from Mexico in 1991.

The sterile male technique also successfully controlled the Mediterranean fruit fly ("medfly"), a destructive fruit fly of citrus, peaches, pears, and apples in California.

Chemical Control

Chemical control refers to the use of pesticides. If chemical control is needed, IPM favors highly targeted chemicals, such as pheromones to disrupt pest mating. **Pheromones** are chemical signals released by animals to communicate with other members of their species. Humans and many insect species alike release pheromones that function in attracting mates. Releasing the pheremones of insect pests can confuse males seeking mates and ultimately hinder them from reproducing (figure 5.3.4.g). This "male confusion" has been successful against the pink bollworm that infests cotton and reduced the need for conventional chemical insecticides by 90%. Pheromones have also been successful against pests that attack tomatoes, grapes, and peaches. If targeted chemicals are not effective, IPM may employ conventional pesticides, ideally only applying them to the spots that they are needed and at the lowest effective concentration. Broadcast spraying of non-specific pesticides is a last resort.



Figure 5.3.4.g: This pheromone trap is used to control the Asiatic rice borer. The caterpillars of this moth species infest rice stems. Image by Mehdi (CC-BY-SA).

Sustainable Practices to Maintain Soil Fertility

A variety of sustainable practices can maintain soil quality. Many of these strategies have additional benefits such as regulating pests, limiting climate change, and preventing water pollution. These methods enrich the soil with nutrients, ensure proper water holding capacity (ability of the soil to retain water), and limit soil-degrading processes such as erosion and compaction.

Crop Rotation

As previously mentioned, **crop rotations** are planned sequences of different crops over time on the same field (figure 5.3.4.h). Rotating crops provides productivity benefits by improving soil nutrient levels and breaking crop pest cycles. Farmers may also choose to rotate crops in order to reduce their production risk through diversification or to manage scarce resources, such as labor, during planting and harvesting. This strategy reduces the pesticide costs by naturally breaking the cycle of weeds, insects and diseases. Also, grass and legumes in a rotation protect water quality by preventing excess nutrients or chemicals from entering water supplies.



Figure 5.3.4.*h*: Crop rotation can be used at a large agricultural scale (left) or in a smaller garden (right). In each situation, which crop species are planted in each location changes year to year. Left image by USDA (public domain) and right image by Sten Porse (CC-BY-SA).

Intercropping

Intercropping means growing two or more crops in close proximity to each other during part or all of their life cycles to promote soil improvement, biodiversity, and pest management. Incorporating intercropping principles into an agricultural operation increases diversity and interaction between plants, arthropods, mammals, birds and microorganisms resulting in a more stable crop-ecosystem and a more efficient use of space, water, sunlight, and nutrients (figure 5.3.4.*i*). This collaborative type of crop management mimics nature and is subject to fewer pest outbreaks, improved nutrient cycling and crop nutrient uptake, and increased water infiltration and moisture retention. Soil quality, water quality and wildlife habitat all benefit.



Figure 5.3.4.*i*: Intercropping of beans and maize (corn). Beans have root nodules that house nitrogen-fixing bacteria, enriching the soil. Image by AnnaJB (CC-BY-SA).

A common example of strip cropping (a type of intercropping; see above) involves alternating a row crop such as corn with a ground-covering crop such as alfalfa. The ground-covering crop helps reduce water runoff and traps soil eroded from the row crop. If this ground-covering crop is a member of the legume family such as alfalfa or soybeans and is associated with nitrogen-fixing bacteria, then alternating the strips from one planting to the next can also help maintain topsoil fertility.

Cover Crops

Cover crops are those that are planted in the off-season to avoid leaving the soil bare. They can prevent soil and wind erosion, improve soil's physical and biological properties, supply nutrients, suppress weeds, improve the availability of soil water, and break pest cycles along with various other benefits. Cover crops are often members of the legume family and help enrich the soil with usable nitrogen. A single species or mix of cover crop species may be planted (figure $5.3.4.j$).



Figure $5.3.4.j$: A roller (crimper) on a tractor being used to roll down a rye and hairy vetch cover crop in April in the United States. While hairy vetch can limit soil erosion and add nitrogen to the soil, it is also an invasive species that must be carefully managed. Image and caption (modified) from Mr. 1032 (CC-BY-SA).

Agroforestry

Agroforestry is the process of planting rows of trees interspersed with a cash crop (figure $5.3.4.k - l$). Besides helping to prevent wind and water erosion of the soil, the trees provide shade which helps promote soil moisture retention. Decaying tree litter also provides some nutrients for the interplanted crops. The trees themselves may provide a cash crop. For example, fruit or nut trees may be planted with a grain crop. You can learn more about agroforestry using this interactive site by the United States Department of Agriculture and Forest Service.



Figure 5.3.4.*k*: Through agroforestry, high-value non-timber crops (food, medicinal plants, woody florals, and crafts) can be cultivated under the protection of a forest canopy that has been managed to provide a favorable crop environment. Image and caption (modified) from USDA/UFS (public domain).

Contour and Terrace Farming

Contour farming involves plowing and planting crop rows along the natural contours of gently sloping land (figure 5.3.4.*l*). The lines of crop rows perpendicular to the slope help to slow water runoff, inhibit the formation of channels of water, and limit soil erosion (and the resultant loss of nutrients). **Terracing** is a common technique used to control water erosion on more steeply sloped hills and mountains (figure 5.3.4.*l*). Broad, level terraces are constructed along the contours of the slopes, and these act as dams trapping water for crops, reducing runoff, and limiting erosion.



Figure 5.3.4.*l*: Contour farming (left) and terracing (right). In contour farming, rows of crops wave perpendicular to the slopes of the land. This example (left) also employs agroforestry. Terrace farming involves making step-like structures (terraces) perpendicular to the incline of a slope to reduce erosion. Left image by Pixture2016 (CC-BY-SA). Right image by USDA (public domain).

Minimal Tillage and No-till Agriculture

In modern agricultural practices, heavy machinery is used to prepare the seedbed for planting, to control weeds, and to harvest the crop. The use of heavy equipment has many advantages in saving time and labor, but can cause compaction of soil and disruption of the natural soil organisms. The problem with soil compaction is that increased soil density limits root penetration depth and may inhibit proper plant growth. Another aspect of soil tillage (mixing the soil) is that it may lead to more rapid decomposition of organic matter due to greater soil aeration. Over large areas of farmland, this has the unintended consequence of releasing more carbon and nitrous oxides (greenhouse gases) into the atmosphere, thereby contributing to climate change.

One of the easiest ways to prevent these problems is to minimize the amount of **tillage**, or turning over of the soil. In **minimal tillage** (conservation tillage) or **no-till agriculture**, the land is disturbed as little as possible by leaving crop residue in the fields (figure 5.3.4.*m*). Special seed drills inject new seeds and fertilizer into the unplowed soil. Tillage of fields does help to break up clods that were previously compacted, so best practices may vary at sites with different soil textures and composition. With proper planning, minimal tillage and no-till agriculture can simultaneously limit soil erosion and compaction, protect soil organisms, reduce costs (if performed correctly), and promote water infiltration. Furthermore, carbon can actually become sequestered into the soil with these methods, thus mitigating climate change. Minimal or no-till agriculture have proved a major success in Latin America and are being used in South Asia and Africa. However, a drawback of this method is that the crop residue can serve as a good habitat for insect pests and plant diseases.



Figure 5.3.4.m: No-till agriculture is an important tool to prevent loss of soil moisture. Image by USDA (public domain).

Windbreaks

Creating **windbreaks** by planting tall trees along the perimeter of farm fields can help control the effects of wind erosion of soil (figure 5.3.4.n). Windbreaks reduce wind speed at ground level, an important factor in wind erosion. They also help trap snow in the winter months, leaving soil less exposed. As a side benefit, windbreaks also provide a habitat for birds and animals. One drawback is that windbreaks can be costly to farmers because they reduce the amount of available cropland.



Figure 5.3.4.n: Windbreaks are rows of trees and shrubs that reduce wind speed. They improve crop yields, reduce soil erosion, improve water-efficiency, protect livestock, and conserve energy. Image and caption (modified) from USDA/UFS (public domain).

Organic Agriculture

Organic agriculture is often incorporated into sustainable agriculture (figure 5.3.4.o). To be certified as organic, farms must avoid using synthetic pesticides, synthetic fertilizers, and genetically modified organisms (figure 5.3.4.p). Organic meat, poultry, eggs, and dairy products come from animals that are given no antibiotics or growth hormones. Pests may instead be controlled by natural enemies or naturally produced substances such as neem oil or diatomaceous earth. Some organic alternatives to synthetic pesticides have relatively low environmental impact (such as salt spray), but there are many naturally produced compounds that are still toxic to people or cause ecological harm when widely used. Interestingly, external application of Bt toxin is approved for organic farming, but use of genetically engineered Bt crops is not. The latter results in lower concentrations of Bt toxin in the environment because it is locally produced directly by the plants themselves. In summary, although many practices of organic agriculture benefit the environment and meet sustainability goals, some organic farms are not sustainable and some sustainable farms are not organic.



Figure 5.3.4.*o*: United States Department of Agriculture (USDA) Conservationist Shemekia Mosley (right) works with Nuri Icgoren (left) who operates Urban Sprout Farms, a biodynamic, certified organic urban farm Lakewood Heights, Georgia. Image and caption (modified) by Preston Keres/USDA (public domain).



Figure 5.3.4.*p*: The USDA certified organic label. Modified from image by nikoretro (CC-BY-SA).

Consumer Choices that Support Sustainable Agriculture

Even if you're not a farmer or lawmaker, you have the power to promote sustainable agriculture as a consumer. As discussed in Food Chains and Food Webs, it generally requires more land and more energy to produce meat compared to plant-based foods due to the inefficient transfer of energy from one trophic level to the next. For this reason, plant-based diets tend to be more sustainable, but this depends on the types of foods consumed and how they were produced.

Local food is not only fresher, but it requires fewer food miles (figure 5.3.4.*q*). In some ways, **organic food** causes less environmental degradation, but note that organic does not necessarily mean sustainable (see above). Because it limits machinery, synthetic pesticides, and synthetic fertilizers, it has a smaller **carbon footprint** (meaning it minimizes contribution to climate change). The Clean 15 is a list of produce that is low in pesticide residues. If you can't afford all organic foods, these are the best items to purchase non-organic. Examples include avocados, sweet corn, and pineapple. The Dirty Dozen lists produce that has the most pesticide residues. If you can only afford a few organic items, these are best to purchase organic. Examples include strawberries, spinach, and nectarines. Finally, some foods, such as beef, have higher **carbon footprint** and **water footprint** than others.



Figure 5.3.4.*q*: Farmers' markets are one option for purchasing locally grown foods. Image by NatalieMaynor (CC-BY).

The Future of the Sustainable Agriculture Concept

Many in the agricultural community have adopted the sense of urgency and direction pointed to by the sustainable agriculture concept (figure 5.3.4.*r*). Sustainability has become an integral component of many government, commercial, and non-profit agriculture research efforts, and it is beginning to be woven into agricultural policy. Increasing numbers of farmers and ranchers have embarked on their own paths to sustainability, incorporating integrated and innovative approaches into their own enterprises.

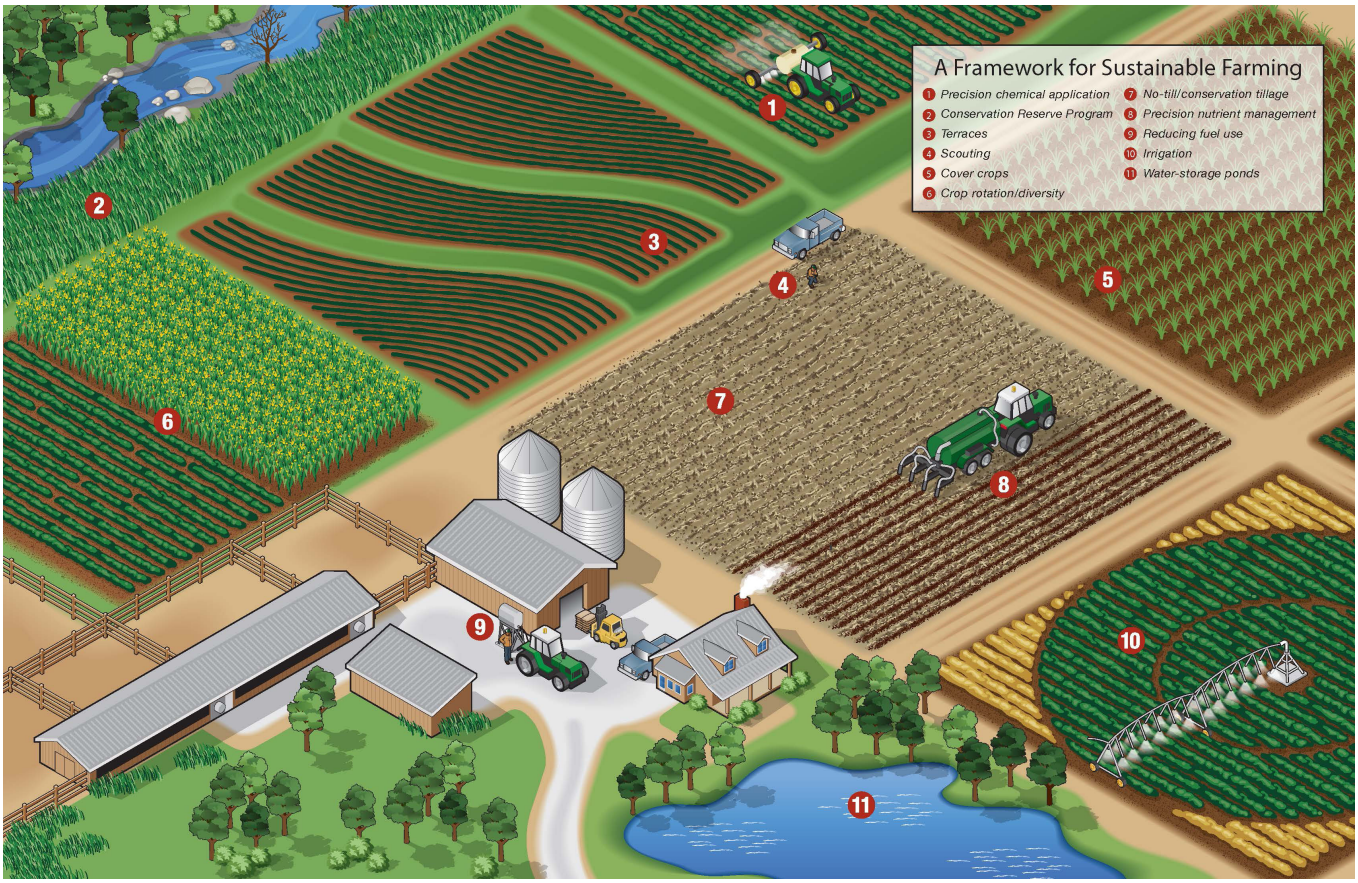


Figure 5.3.4.r: Sustainable agriculture combines many approaches. (1) If chemical are used, they are applied only where needed. (2) Through the Conservation Reserve Program, the most ecologically important patches are left intact. (3) Terraces on slopes reduce soil erosion. (4) Crops are scouted for pests as part of monitoring, a component of integrated pest management. (5) Cover crops prevent erosion and enrich the soil during off-seasons. (6) Different types of crops are included are planted in close proximity (intercropping) or in different years (crop rotation). (7) Land is either not tilled at all or tilled minimally (conservation tillage). (8) Through precision nutrient management, soil is sampled at different locations, and fertilizer is only applied where needed. (9) Heavy equipment is used only as needed and in such a way that conserves fuel. (10) Irrigation is precisely timed to conserve water. (11) Water storage ponds help recharge groundwater. Image by the United Soybean Board (CC-BY).

Attributions

Modified by Melissa Ha from the following sources:

- Sustainable Agriculture from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Soils from *AP Environmental Science* by University of California College Prep, University of California (licensed under CC-BY). Download for free at CNX.
- Sustainability Definitions. USDA. Accessed 01-05-2021. (public domain)
- Biological Control from *Biology* by John W. Kimball (licensed under CC-BY)
- Integrated Pest Management (IPM) Principles. EPA. Accessed 01-05-2021 (public domain)
- Cover Crops and Soil Health. Natural Resources Conservation Service, USDA. Accessed 01-06-2021. (public domain)

This page titled 5.3.4: Sustainable Agriculture is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **14.4: Sustainable Agriculture** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

5.3.5: Data Dive- Landslides and Ag Terraces

Overview

The beauty of terraced agriculture almost makes the observer forget that it is still a form of habitat destruction and landscape modification. A 2019 study investigated the risk to terraced landscapes with varying states of upkeep after severe rainfall events, which increase chances of serious damage due to landslide events. The authors compared three classes of upkeep:

- Succession on terrace: Which includes all terraced areas in which cultivation activities ceased, and is currently substituted by both wood and shrub cover.
- Cultivated terrace: Which includes all areas which are still under any kind of cultivation, from olive groves, to vineyards, to other crops
- Vegetated natural slopes: Which refer to those areas covered by spontaneous vegetation which have never been terraced.

Landslide mapping was used to estimate the size of damage after a landslide event took place. Below is a graph for some of the published results:

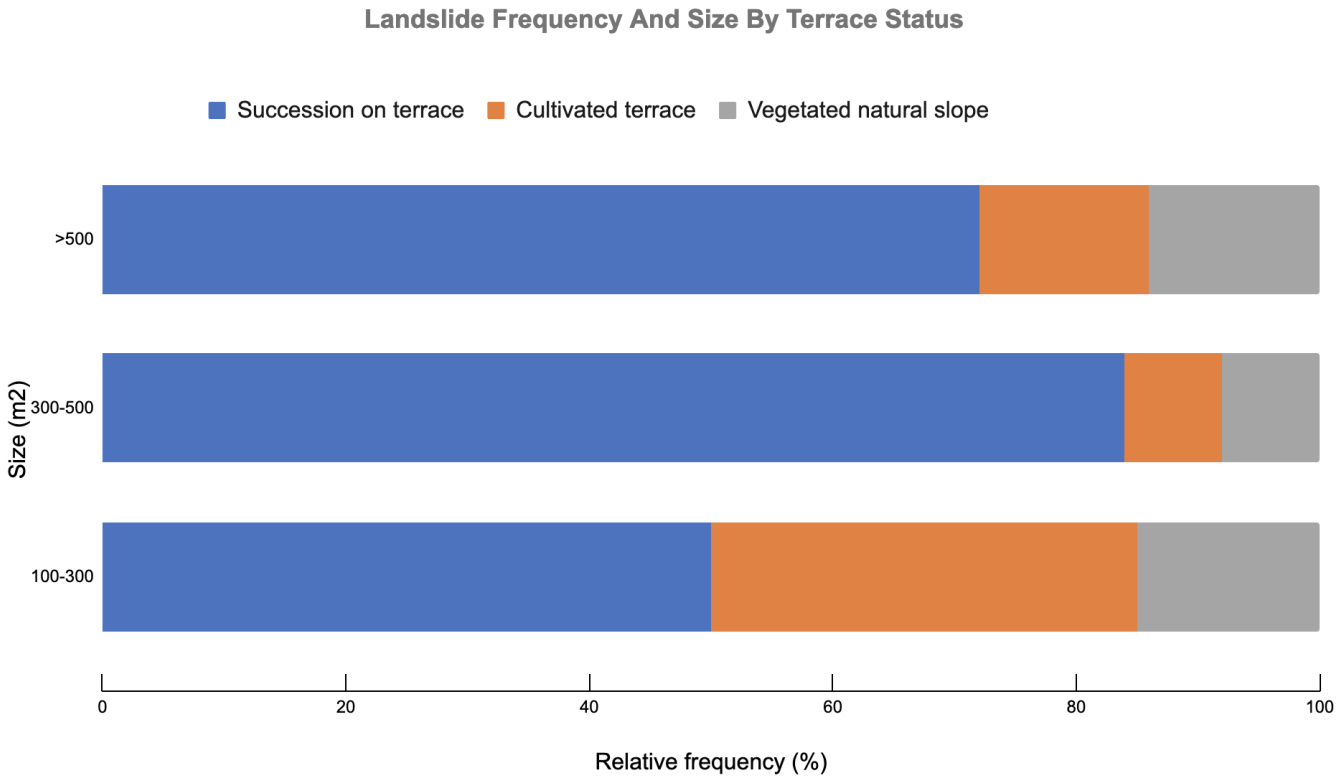


Figure 5.3.5.*a*: Percent breakdown of landslide frequency and size by terrace status. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Agnoletti M, Errico A, Santoro A, Dani A, Preti F 2019.

Questions

1. What is the independent (explanatory) variable and the dependent (response) variable?
2. What question(s) are the authors trying to answer with this graph?
3. Do the authors use any control category/variable to help clarify the patterns observed? If so, what?
4. Which status has the most consistent increased landslide frequency?
5. What landslide frequency patterns can you observe across the three size?
6. How can the results of this graph to inform future projects/restoration efforts when a terrace is abandoned?

Raw Data For Above Graph(s)

Table 5.3.5.*a*: Raw data for percent breakdown of landslide frequency and size by terrace status. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Agnoletti M, Errico A, Santoro A, Dani A, Preti F 2019.

Terrace Status	100-300	300-500	Over 500
Succession On Terrace	50	84	72
Cultivated Terrace	35	8	14
Vegetated Natural Slope	15	8	14

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 5.3.5: Data Dive- Landslides and Ag Terraces is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **14.5: Data Dive- Landslides and Ag Terraces** by Melissa Ha and Rachel Schleiger has no license indicated.

5.3.6: Review

Summary

After completing this chapter you should be able to...

- Describe the global state of food security.
- Compare industrial agriculture, sustainable agriculture, and organic agriculture.
- Define pests and pesticides.
- Detail the advantages and disadvantages of industrial agriculture.
- Compare selective breeding with genetic engineering.
- Explain how plants are genetically modified.
- Detail the advantages and disadvantages of using genetically modified crops.
- Explain the steps of integrated pest management and provide examples of cultural, mechanical, biological, and chemical control.
- Describe each sustainable practice that maintains soil quality and explain the additional benefits of these practices.

Progress continues in the fight against hunger, yet an unacceptably large number of people still lack the food they need for an active and healthy life. Food security is dependent on availability, access, and utilization.

Industrial agriculture is the prevailing agricultural system that relies on heavy machinery, synthetic pesticides, and synthetic herbicides. It has led to tremendous gains in productivity and efficiency worldwide. On the other hand, industrial agriculture has put small farms at a disadvantage and caused several ecological issues. It has exacerbated climate change, caused water pollution, degraded soil quality, harmed beneficial native organisms, and led to the evolution of pesticide resistance.

Both selective breeding and genetic engineering are processes of agricultural biotechnology. **Selective breeding** occurs when humans assist individual organisms with favorable traits in reproduction generation after generation. **Genetic engineering** refers to lab methods that directly alter an organism's DNA. Advocates say that application of genetic engineering in agriculture has resulted in benefits to farmers, producers, and consumers. While genetically modified organisms (GMOs) have potential to address malnutrition, reduce pesticide application, and combat crop disease, the companies who patent them now have greater control over agriculture. Furthermore, escaped transgenes can create herbicide-resistant superweeds. While members of the public have raised concerns about health impacts of GMOs, the National Academies of Sciences, Engineering, and Medicine's assessment of over 900 studies found that GMOs do not threaten human health.

Sustainable agriculture is an alternative to industrial agriculture that focuses on preserving natural resources and protecting the environment. It can thus be continued indefinitely. Sustainable agriculture employs **integrated pest management** (IPM), which relies on a variety of approaches to keep populations of true pests below economically harmful levels. Approaches with the lowest environmental impact are prioritized under IPM. A variety of sustainable practices such as crop rotation, intercropping, terrace farming, and minimal tillage can maintain soil fertility. Many of these methods additionally help regulate pest populations, preserve water quality, and combat climate change. **Organic agriculture** avoids using synthetic pesticides, synthetic fertilizers, GMOs, hormones, and antibiotics. Organic and sustainable farming practices often overlap. Consumers can promote sustainable agriculture by purchasing local, organic foods and choosing foods with low carbon and water footprints.

Attribution

Modified by Melissa Ha from Conventional and Sustainable Agriculture and Food and Hunger from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 5.3.6: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **14.6: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

5.4: Environmental Health

Chapter Hook

In 1854 England, a cholera epidemic started wiping out people within a day of symptoms starting. Only one physician, Dr John Snow, thought that it connected with contaminated water sources. Over the course of a month, Dr Snow was able to provide evidence that one particular well was connected with almost all of the deaths. Later, it was discovered that the contaminant in the water was a bacterium, named for the disease it caused, *Vibrio cholerae*. Unfortunately, cholera is not a disease of the past. The World Health Organization calculates an average of 1.3-4 million cases where an estimated 21-143,000 of these cases result in death. In some places, there is enough preventative testing to stop the spread of the disease before people get sick. However, in others, the contamination is unfortunately addressed only after an epidemic occurs.

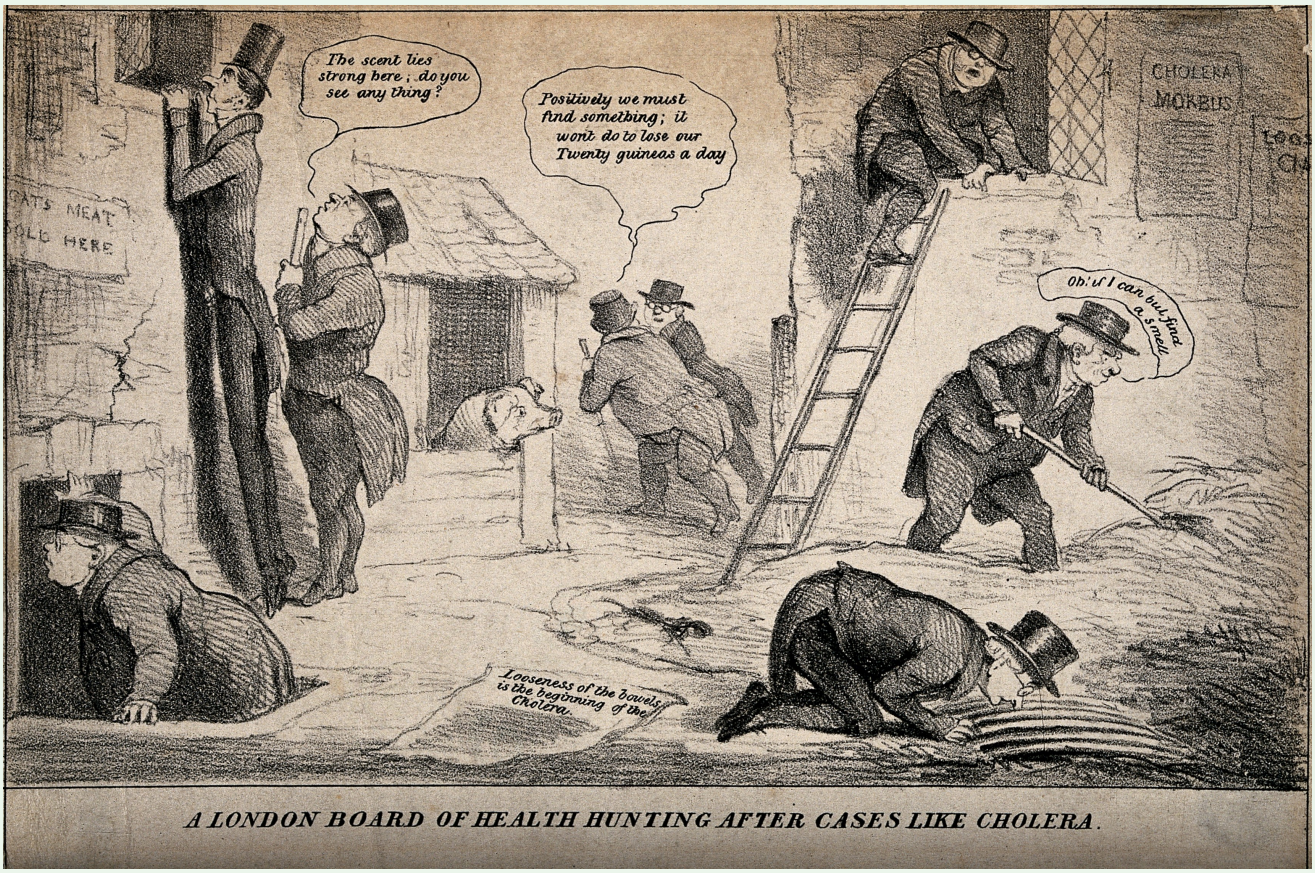


Figure 5.4.*a*: Cholera cartoon from 1854. Image by Wikimedia Commons (CC-BY-2.0)

5.4.1: Types of Environmental Hazards

5.4.2: Epidemiology

5.4.3: Infectious Diseases

5.4.4: Environmental Toxicology

5.4.5: Environmental Hazard Reduction

5.4.6: Data Dive- Cholera Cases Worldwide

5.4.7: Review

Attribution

Modified by Rachel Schleiger (CC-BY-NC).

This page titled 5.4: Environmental Health is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

5.4.1: Types of Environmental Hazards

Environmental health is a field that focuses on how the natural and human-built surroundings as well as behaviors affect human well-being. The field is concerned with preventing disease, death, and disability by reducing exposure to environmental hazards and promoting behavioral change. **Environmental hazards** are threats to human health and well-being (table 5.4.1.a).

Table 5.4.1.a: Typical Environmental Health Issues: Determinants and Health Consequences.

Underlying Determinants	Possible Adverse Health and Safety Consequences
Inadequate water (quantity and quality), sanitation and solid waste disposal, improper hygein (handwashing)	Diarrhea and vector-related diseases (for example, malaria, schistosomiasis, and dengue)
Improper water resource managment, uncluding poor drainage	Vector-related diseases
Crowded housing and poor ventialtion of smoke	Acute and chronic respiratory disease, including lung cancer from coal and tobacco inhalation
Exposures to vehicular and industrial air pollution	Respiratory diseases, some cancers, and loss of IQ in children
Population movement and encroachment and construction, which affect feeding and breeding grounds of vectors, such as mosquitoes	Vector-related diseases May also spread other infectious diseases (for example, HIV/AIDS, Ebola)
Exposure to naturally ocurring toxic substances	Poising from substances such as arsenic, manganese, and fluorides
Natural resources degradation (for example, landslides, poor drainage, erosion)	Injury and death from landslides and flooding
Climate change, partly from combustion of fossil fuel and release of greenhouse gases in transportation, industry, and poor energy conservation in housing, fuel, commerce, and industry	Injury/death from extreeme heat/cold, sotres, floods, and fires Indirect effects spread of vectorbrone diseases, aggravation of respriatory diseases, population dislocation, water pollution from sea level rise, etc.
Ozone depletion from industrial and commercial actvitiy	Skin cancer, cataracts Indirect effect: compromised food production

Table based on Lvovsky/World Bank (CC-BY)

Traditional versus Modern Environmental Hazards

Environmental hazards can be classified as traditional or modern. **Traditional hazards** are related to poverty and mostly affect low-income people and those in developing countries. **Modern hazards**, caused by technological development, prevail in industrialized countries where exposure to traditional hazards is low.

The impact of traditional hazards exceeds that of modern hazards by 10 times in Africa, five times in Asian countries (except for China), and 2.5 times in Latin America and the Middle East (figure 5.4.1.a). Water-related diseases caused by inadequate water supply and sanitation impose an especially large health burden in Africa, Asia, and the Pacific region. In India alone, 90,000 children under five died diarrhea in 2017. Globally, 409,000 people died of malaria in 2019 with 94% of these deaths occurring in African countries. In 2016, approximately one third of the world's households used unprocessed solid fuels, particularly biomass (crop residues, wood, and dung) for cooking and heating in inefficient stoves without proper ventilation. This exposes people—mainly low-income women and children—to high levels of indoor air pollution, the cause of about 1.6 million deaths in each year (figure 5.4.1.b).

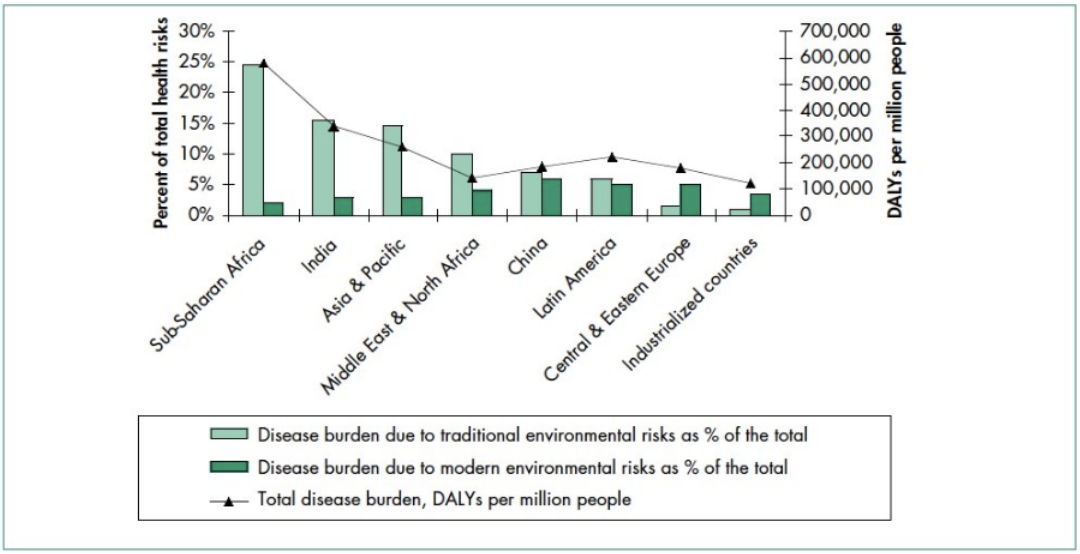


Figure 5.4.1.a: The percent of total health risks due to traditional (light green bars on the left) versus modern (dark green bars on the right) health risks and total disease burden (line and triangles) in disability-adjusted life years (DALY) per million people in eight regions. The disability-adjusted life years measures the burden of disability associated with a disease or disorder and represent represent the total number of years lost to illness, disability, or premature death within a given population. Traditional environmental health hazards prevail in developing countries, but modern risks are also significant. Traditional hazards are greatest in Sub-Saharan Africa (25%), and modern hazards are greatest in China (5%), Latin America (4%), and Central and Eastern Europe (4%). For most regions, traditional hazards were a greater threat than modern hazards with the exceptions of Central and Eastern Europe and industrialized countries. Image from World Bank (CC-BY).

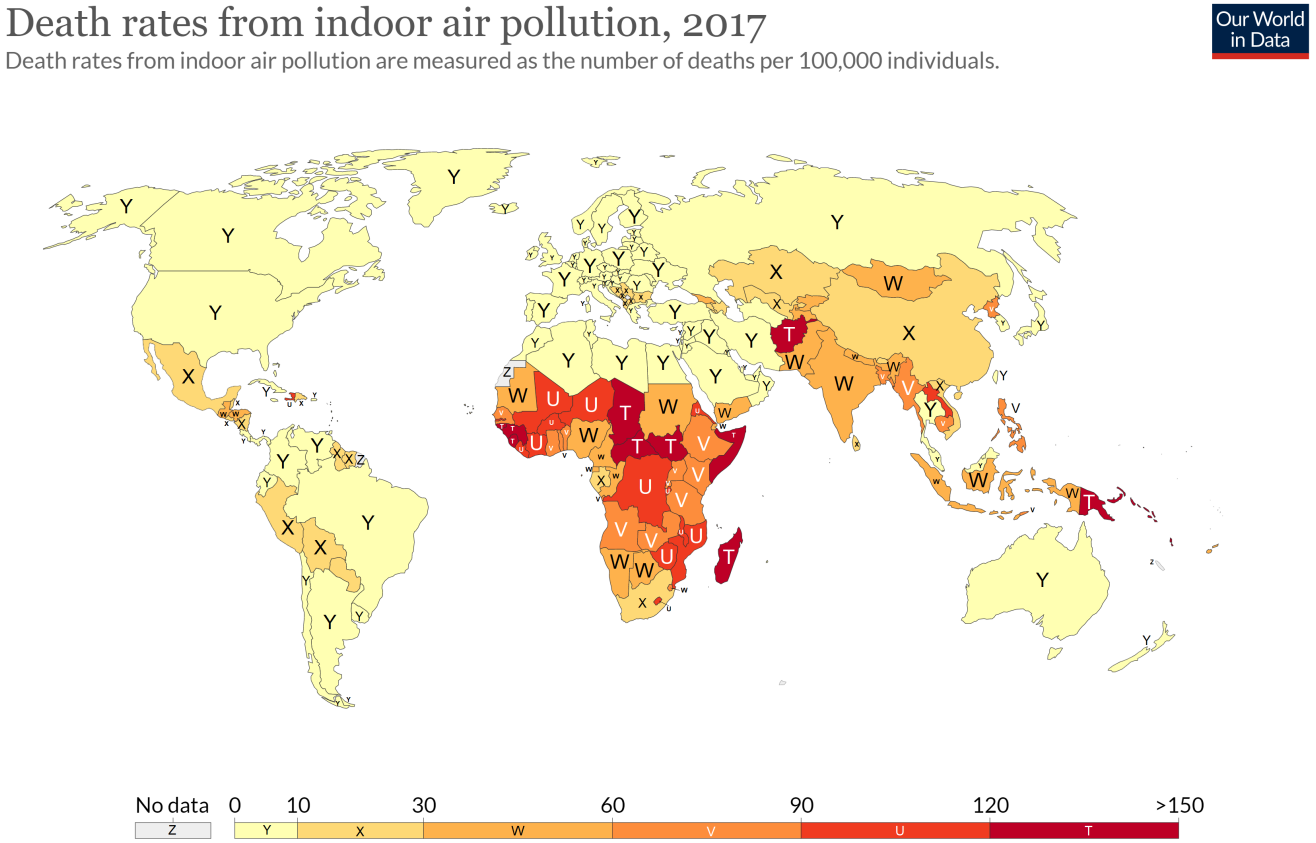


Figure 5.4.1.*b*: Death rates from indoor air pollution in each country in 2017. Guinea, Sierra Leone, Chad, Central African Republic, South Sudan, Somalia, Madagascar, Afghanistan, and Papua New Guinea had the highest death rates (more than 120 deaths per 100,000 individuals, T). Moderately high death rates (90-120 deaths per 100,000 individuals, U) were mostly in African and Southeast Asian countries. Moderate death rates (30-90 deaths per 100,000 individuals, W and V) were mostly in African and Asian countries. Moderately low death rates (10-30 deaths per 100,000 individuals, X) were in Central and South America and northern Asia. The lowest death rates were (0-10 deaths per 100,000 individuals, Y) were in mostly in North American, South American, European, and North African countries and Australia. Image modified from Hannah Ritchie (2013). Indoor Air Pollution. Published online at OurWorldInData.org. (CC-BY).

The contribution of modern environmental risks to the disease burden in most developing countries is similar to – and in quite a few countries, greater than – that in rich countries (figure 5.4.1.*a*). Urban air pollution, for example, is highest in parts of China, India, and some cities in Asia and Latin America. Low-income people increasingly experience a "**double burden**" of traditional and modern environmental health risks. In rich countries, they experience twice the burden of illness and death from all causes and 10 times greater disease burden from environmental risks.

Biological, Chemical, and Physical Environmental Hazards

Environmental hazards can also be classified into three interrelated categories (biological, chemical, and physical) based on the properties of their causes. These categories are not mutually exclusive with traditional versus modern hazards. For example, indoor air pollution is both a traditional and chemical hazard. Different hazards can interact and exacerbate one another. For example, a flood is primarily a physical hazard, but it can lead to the spread of waterborne disease (a biological hazard). Similarly, air pollution (a chemical hazard) can damage respiratory tissue, making the body more vulnerable to a respiratory infection (a biological hazard). Infectious diseases (biological hazards) can also weaken the immune system, making an individual more vulnerable to chemical hazards.

Biological Hazards

For most of human history, biological hazards were the most significant factor in health. **Biological hazards** are infectious (communicable) diseases caused by **pathogens** (disease-causing organisms or infectious particles) such as bacteria, fungi, parasitic worms, protozoa, viruses, and prions. **Bacteria** are single-celled organisms with small, simple cells. Examples of bacterial diseases include tuberculosis, cholera, bacterial pneumonia, and dysentery. **Fungi** may have one or multiple cells and have a more complex cell type than bacteria. Fungal diseases include minor infections like candidiasis (yeast infection) or athlete's foot, but they can also causes severe respiratory infections (histoplasmosis, coccidioidomycosis, etc.) particularly in individuals with compromised immune systems. **Parasitic worms** are animals from several phyla (groups) that siphon nutrients from their hosts. Examples include tapeworms, commonly acquired through consuming undercooked meat, and blood flukes (*Schistosoma*). Like fungi, **protozoa** have larger, more complex cells than bacteria, but they are single celled and lack the rigid cell wall that surrounds fungal cells. Malaria (figure 5.4.1.*c*), African trypanosomiasis (sleeping sickness), and giardiasis are caused by protozoa. **Viruses** are infectious particles with genetic information surrounded by a protein coat, but they are not technically considered organisms in part because they do not consist fo cells. COVID-19, influenza, measles, the common cold, ebola viral disease (Ebola hemorrhagic fever), and human immunodeficiency virus (HIV)/acquired immune deficiency syndrome (AIDS) are all caused by viruses. **Prions** (proteinaceous infectious particles) are even simpler than viruses because they lack genetic material and only contain protein.

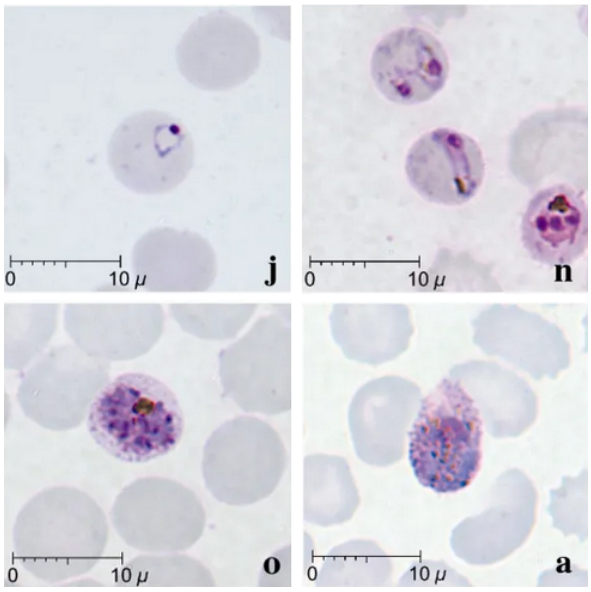


Figure 5.4.1.*c*: Red blood cells infected with the protozoan *Plasmodium*, which causes malaria, under the microscope. Image by Kim-Sung Lee, Janet Cox-Singh, Balbir Singh (CC-BY).

While the proportion of deaths caused by in infectious diseases has overall decreased (with a higher proportion of deaths caused by noncommunicable diseases such as cancer and cardiovascular disease), infectious diseases still caused about one in five deaths in 2017. These deaths occurred at the highest rates in developing countries and many were in children. Malnutrition, unclean water, poor sanitary conditions and lack of proper medical care all play roles in transmission and high death rates from infectious diseases Compounding the problems of infectious diseases are factors such as antibiotic-resistant pathogens, pesticide-resistant disease vectors, and overpopulation.

Chemical Hazards

Chemical hazards are toxic substances, which cause damage to living organisms. Air pollutants (such as secondhand smoke or carbon monoxide), heavy metals, and pesticides are a few examples. We can be exposed to these contaminants from a variety of residential, commercial, and industrial sources. Sometimes harmful environmental contaminants occur biologically, such as those from mold or a toxic algae bloom. Toxins can be classified based on their origin, purpose, chemical structure and properties, or effects. Table 5.4.1.*b* describes a few categories of toxins based on their effects and provides examples. A few of these examples are discussed more specifically below.

Table 5.4.1.*b*: Classification of Environmental Contaminants

Contaminant	Definition
Carcinogen	An agent which may produce cancer (uncontrolled cell growth), either by itself or in conjunction with another substance. Examples include formaldehyde, asbestos, radon, vinyl chloride, and tobacco.
Teratogen	A substance which can cause physical defects in a developing embryo. Examples include alcohol and cigarette smoke.
Mutagen	A material that induces genetic changes (mutations) in the DNA. Examples include radioactive substances (such as radon and nuclear fuel and waste) and nitrous acid. Some forms of radiation (see Physical Hazards) are also mutagens.
Neurotoxin	A substance that can cause an adverse effect on the chemistry, structure or function of the nervous system. Examples include lead and mercury.
Endocrine disruptor	A chemical that may interfere with the body's endocrine (hormonal) system and produce adverse developmental, reproductive, neurological, and immune effects in both humans and wildlife. A wide range of substances, both natural and man-made, are thought to cause endocrine disruption, including pharmaceuticals, dioxin and dioxin-like compounds, arsenic, polychlorinated biphenyls (PCBs), DDT and other pesticides, per- and polyfluoroalkyl substances (PFAS), pthalates, and plasticizers such as bisphenol A (BPA).

Formaldehyde

Formaldehyde is a colorless, flammable gas or liquid that has a pungent, suffocating odor. It is a **volatile organic compound**, which is a compound containing carbon and hydrogen that easily becomes a vapor or gas. It is also naturally produced in small, harmless amounts in the human body. The primary way we can be exposed to formaldehyde is by breathing air containing it. Formaldehyde is released into the air by industries using or manufacturing formaldehyde, wood products (such as particle-board, plywood, and furniture), automobile exhaust, cigarette smoke, paints and varnishes, and carpets and permanent press fabrics. Nail polish and commercially applied floor finish emit formaldehyde (figure 5.4.1.*d*).



Figure 5.4.1.*d*: Nail products are known to contain toxic chemicals, such as dibutyl phthalate (DBP), toluene, and formaldehyde.

In general, indoor environments consistently have higher concentrations than outdoor environments because many building materials, consumer products, and fabrics emit formaldehyde. Levels of formaldehyde measured in indoor air range from 0.02–4 parts per million (ppm). Formaldehyde levels in outdoor air range from 0.001 to 0.02 ppm in urban areas.

Heavy Metals

Heavy metals are chemical elements of high density that form a special type of bond (called metallic bonds, in which electrons are shared but in a less constricted way than in covalent bonds). Arsenic, mercury, lead, and cadmium are examples of heavy metals.

Arsenic (As) is a naturally occurring element that is normally present throughout our environment in water, soil, dust, air, and food. Levels of arsenic can regionally vary due to farming and industrial activity as well as natural geological processes. The arsenic from farming and smelting tends to bind strongly to soil and is expected to remain near the surface of the land for hundreds of years as a long-term source of exposure. Wood that has been treated with chromated copper arsenate (CCA) is commonly found in decks and railings in existing homes and outdoor structures such as playground equipment. Some underground aquifers are located in rock or soil that has naturally high arsenic content.

Most arsenic gets into the body through ingestion of food or water. Arsenic in drinking water is a problem in many countries around the world, including Bangladesh, Chile, China, Vietnam, Taiwan, India, and the United States. Arsenic may also be found in foods, including rice and some fish, where it is present due to uptake from soil and water. It can also enter the body by breathing dust containing arsenic.

Arsenic poisoning causes a variety of symptoms and serious health conditions (figure 5.4.1.*e*). Researchers are finding that arsenic, even at low levels, can interfere with the body's endocrine system. Arsenic is also a known human carcinogen associated with skin, lung, bladder, kidney, and liver cancer.



Figure 5.4.1.*e*: Patchy areas of dark skin pigmentation (arsenical hyperkaratosis) on the palms of the hands is a symptom of arsenic poisoning. Image and caption (modified) from Agency for Toxic Substances and Disease Registry/CDC (public domain).

Mercury (Hg) is a naturally occurring metal, a useful chemical in some products, and a potential health risk. Mercury exists in several forms; the types people are usually exposed to are methylmercury and elemental mercury. Elemental mercury at room temperature is a shiny, silver-white liquid which can produce a harmful odorless vapor. Methylmercury, an organic compound, can build up in the bodies of long-living, predatory fish (Biomagnification). Although fish and shellfish have many nutritional benefits, consuming large quantities of fish increases a person's exposure to mercury. Pregnant women who eat fish high in mercury on a regular basis run the risk of permanently damaging their developing fetuses. Children born to these mothers may exhibit motor difficulties, sensory problems and cognitive deficits. The United States Environmental Protection Agency thus recommends that pregnant women and young children should not consume any swordfish, shark, king mackerel, or tilefish because of their high mercury content. These individuals are advised to eat fish low in mercury such as salmon, shrimp, pollock, and catfish (figure 5.4.1.*f*). To keep mercury out of the fish we eat and the air we breathe, it's important to take mercury-containing products to a hazardous waste facility for disposal. Common products sold today that contain small amounts of mercury include fluorescent lights and button-cell batteries (figure 5.4.1.*g*).



Figure 5.4.1.f: Fish classified based on mercury levels. The best choices (black sea bass, catfish, herring, trout, and many others) have the lowest mercury levels, and two or three servings of these choices can be safely consumed each week. The good choices (carp, halibut, yellowfin tuna, etc.) have moderate levels, and it is safe to eat one serving per week. The choices to avoid, such as shark and swordfish, have the highest mercury levels and should be avoided. A table listing all fish in each category is available here. Image by EPA and FDA (public domain).



Figure 5.4.1.g: Button-cell batteries found in small devices like watches and hearing aids contain mercury and must be discarded at the proper hazardous waste facility. Image by Lead holder (CC-BY-SA).

Lead (Pb) is a metal that occurs naturally in the rocks and soil of the Earth's crust. It is also released from mining, manufacturing, and the combustion (burning) fossil fuels such as coal, oil, gasoline, and natural gas. Lead has no distinctive taste or smell. Lead is used to produce batteries, pipes, roofing, scientific electronic equipment, military tracking systems, medical devices, and products to shield X-rays and nuclear radiation. It is used in ceramic glazes and crystal glassware. Because of health concerns, lead and lead compounds were banned from house paint in 1978; from solder used on water pipes in 1986; from gasoline in 1995; from solder used on food cans in 1996; and from tin-coated foil on wine bottles in 1996. The U.S. Food and Drug Administration has set a limit on the amount of lead that can be used in ceramics.

Lead and lead compounds are listed as "reasonably anticipated to be a human carcinogen". It can affect almost every organ and system in your body. It can be equally harmful if breathed or swallowed. The part of the body most sensitive to lead exposure is the central nervous system, especially in children, who are more vulnerable to lead poisoning than adults. A child who swallows large amounts of lead can develop brain damage that can cause convulsions and death; the child can also develop blood anemia, kidney damage, colic, and muscle weakness. Repeated low levels of exposure to lead can alter a child's normal mental and physical growth and result in learning or behavioral problems. Exposure to high levels of lead for pregnant women can cause miscarriage,

premature births, and smaller babies. Repeated or chronic exposure can cause lead to accumulate in your body, leading to lead poisoning.

The video below explains how Flint, Michigan's water supply was polluted with lead in 2014 and the problematic government response that followed. This video was made in 2016. As of 2020, Flint has a clean water source, and the city of Flint is in the process of compensating affected residents for the damages. (Click here to read a 2020 update.)



Asbestos

Asbestos is a mineral fiber that occurs in rock and soil. Because of its fiber strength and heat resistance asbestos has been used in a variety of building construction materials for insulation and as a fire retardant. Asbestos has also been used in a wide range of manufactured goods, mostly in building materials (roofing shingles, ceiling and floor tiles, paper products, and asbestos cement products), friction products (automobile clutch, brake, and transmission parts), heat-resistant fabrics, packaging, gaskets, and coatings. Exposure to asbestos is associated with cancers (lung cancer and mesothelioma) and another lung disease called asbestosis. In the United States, certain uses of asbestos, including in corrugated paper, flooring, and building insulation, are banned under the Toxic Substances Control Act and Clean Air Act (figure 5.4.1.h). In contrast, asbestos is fully banned in 67 countries as of 2019.



Figure 5.4.1.h: An asbestos pipe wrap. While this use of asbestos is banned in the United States, other uses are still permitted. Image by EPA (public domain).

Per- and polyfluoroalkyl substances (PFAS)

Per- and polyfluoroalkyl substances (PFAS) are a group of manufactured organic chemicals used in a variety of industries (figure 5.4.1.i). They can be found in food packaging, stain- and water-repellent fabrics, nonstick products (such as Teflon), polishes,

waxes, paints, cleaning products, and fire-fighting foams.

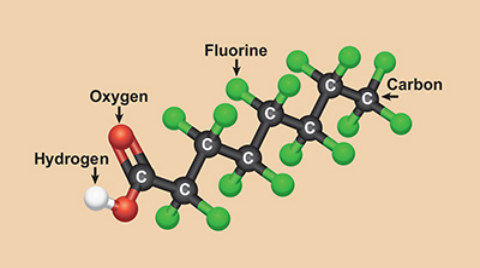


Figure 5.4.1.*i*: The chemical structure of perfluorooctanoic acid (PFOA), a PFAS. This molecule has eight carbon atoms (black) connected in a chain. Most of them are bound to fluorine atoms (green). The carbon at the end of the chain forms two covalent bonds with an oxygen and is single bounded to another oxygen atom, which is also attached to a hydrogen. Image by National Institute of Environmental Health Sciences (public domain).

Studies indicate that some PFAS can cause reproductive and developmental, liver and kidney, and immunological effects in laboratory animals. More limited findings associate some PFAS with low infant birth weights, effects on the immune system, cancer, and thyroid hormone disruption in humans.

Eight major chemical manufacturers in the U.S. phased out the use certain PFAS (called perfluorooctanoic acid, PFOA, and perfluorooctane sulfonate, PFOS) and related chemicals in their products and as emissions from their facilities. However, these PFAS can still be imported and other PFAS are still manufactured in the U.S.

Polychlorinated Biphenyls (PCBs)

Polychlorinated biphenyls (PCBs) are a group of manufactured organic chemicals. They belong to a broad family of chemicals known as chlorinated hydrocarbons, which consisting of carbon, hydrogen and chlorine atoms (figure 5.4.1.*j*). The number of chlorine atoms and their location in a PCB molecule determine many of its physical and chemical properties. PCBs have no known taste or smell, and range in consistency from an oil to a waxy solid.

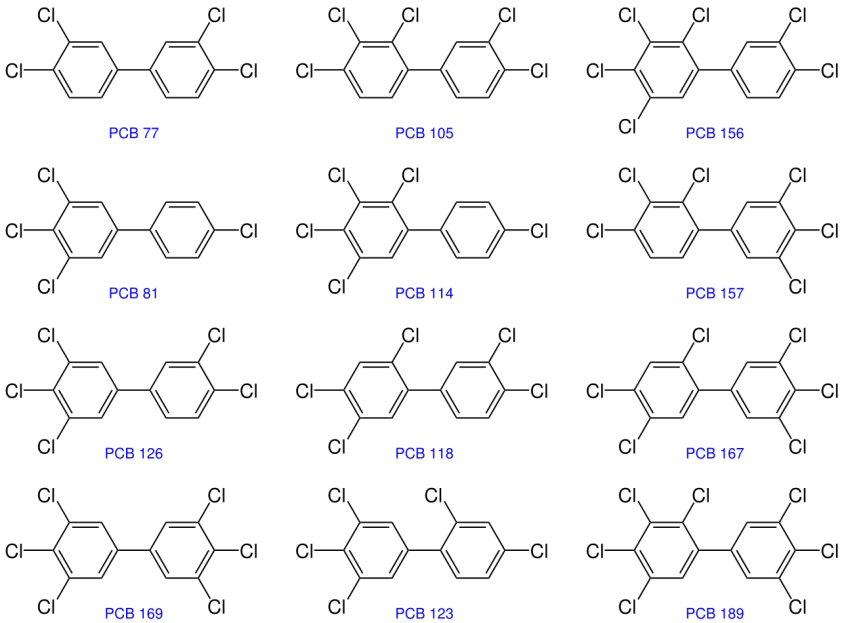


Figure 5.4.1.*j*: A variety of polychlorinated biphenyls (PCBs). They each have two rings of six carbon atoms each attached together. Additionally, each has multiple chlorine atoms attached to the rings, but the exact number and placement of chlorine atoms varies. Image by Leyo/M. Van den Berg et al. (public domain).

Polychlorinated biphenyls have been shown to cause cancer, cause birth defects, and affect the immune, reproductive, nervous, and endocrine systems in animals. Studies in humans support evidence for potential carcinogenic and non-carcinogenic effects of

PCBs. The different health effects of PCBs may be interrelated. Alterations in one system may have significant implications for the other systems of the body.

The manufacture of PCBs in the U.S. began in 1929 until it was banned in 1979 under the Toxic Substances Control Act. Due to their non-flammability, chemical stability, high boiling point and electrical insulating properties, PCBs were used in hundreds of industrial and commercial applications including in electrical equipment, coolants paints, plastics, rubber products, pigments, and dyes.

Bisphenol A (BPA)

Bisphenol A (BPA) is a chemical synthesized in large quantities for use primarily in the production of polycarbonate plastics and epoxy resins. Polycarbonate plastics have many applications including use in some food and drink packaging such as water and infant bottles, impact-resistant safety equipment, and medical devices (figure 5.4.1.*k*). Epoxy resins are used as lacquers to coat metal products such as food cans, bottle tops, and water supply pipes. Some dental sealants and composites may also contribute to BPA exposure. The primary source of exposure to BPA for most people is through the diet. Bisphenol A can leach into food from the protective internal epoxy resin coatings of canned foods and from consumer products such as polycarbonate tableware, food storage containers, water bottles, and baby bottles. The degree to which BPA leaches from polycarbonate bottles into liquid may depend more on the temperature of the liquid or bottle, than the age of the container. It has also be found in breast milk.



Figure 5.4.1.*k*: A water bottle is labeled as BPA free. Image by Hteink.min (CC-BY-SA).

Some animal studies suggest that infants and children may be the most vulnerable to the effects of BPA. It disrupts signaling by estrogen, a naturally produced hormone, and the U.S. National Toxicology Program (NTP) documented concerns about its effects on the behavior, brain, and prostate in young children and developing fetuses.

The following personal choices can reduce exposure to BPA:

- Avoid microwaving polycarbonate plastic food containers. Polycarbonate is strong and durable, but over time it may break down from over use at high temperatures.
- Note recycle codes on the bottom of plastic containers. Some, but not all, plastics that are marked with recycle codes 3 or 7 may be made with BPA.
- Reduce your use of canned foods.
- When possible, opt for glass, porcelain or stainless steel containers, particularly for hot food or liquids.

While BPA is not banned in the U.S., the Food and Drug Adminsitration banned its use in baby bottles and sippy cups in 2012 and its use in the coating of infant formula containers in 2013. However, similar compounds such as bisphenol S (BPS) are now used as replacements.

Phthalates

Phthalates are a group of synthetic chemicals used to soften and increase the flexibility of plastic and vinyl. Polyvinyl chloride is made softer and more flexible by the addition of phthalates. Phthalates are used in hundreds of consumer products. Phthalates are used in cosmetics and personal care products, including perfume, hair spray, soap, shampoo, nail polish, and skin moisturizers (figure 5.4.1.*l*). They are used in consumer products such as flexible plastic and vinyl toys, shower curtains, wallpaper, vinyl miniblinds, food packaging, and plastic wrap. Exposure to low levels of phthalates may come from eating food packaged in plastic

that contains phthalates or breathing dust in rooms with vinyl miniblinds, wallpaper, or recently installed flooring that contain phthalates. We can be exposed to phthalates by drinking water that contains phthalates.



Figure 5.4.1.*l*: Phthalates are often used in shampoos and other personal care products, but some brands produce phthalate-free products. Image by inf per Open Food Facts (CC-BY-SA).

Phthalates are suspected to be endocrine disruptors. Some types of phthalates have affected the reproductive system of laboratory animals. In 2017, The U.S. Consumer Product Safety Commission (CPSC) has banned several pthalates from being used at concentrations greater than 0.1% in toys and products designed to be used by children three years old or younger.

Radon

Radon is a radioactive gas that is naturally-occurring, colorless, and odorless (figure 5.4.1.*m*). It comes from the natural decay of uranium or thorium found in nearly all soils. It typically moves up through the ground and into the home through cracks in floors, walls and foundations. It can also be released from building materials or from well water. Radon breaks down quickly, giving off radioactive particles. Long-term exposure to these particles can lead to lung cancer. Radon is the leading cause of lung cancer among nonsmokers, according to the U.S. Environmental Protection Agency, and the second leading cause behind smoking. To reduce the risk of radon exposure, the Department of Urban and Housing Development recommends testing your home for radon, avoiding smoking to reduce the risk of lung cancer, and ensuring proper ventilation in your home.

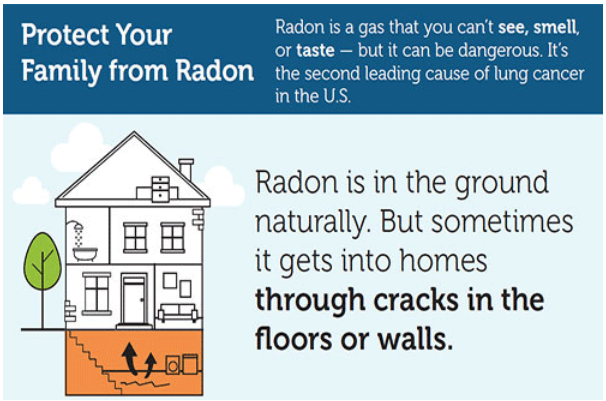


Figure 5.4.1.*m*: This information graphic from the Centers for Disease Control explains how people can be exposed to radon from underground and its health risks. It reads "Protect your family from radon. Radon is a gas that you can't see, smell, or taste— but it can be dangerous. It's the second leading cause of lung cancer in the U.S. Radon is in the ground naturally. But sometimes it gets into homes through cracks in the floors or walls." Image by CDC (public domain).

Dichlorodiphenyltrichloroethane (DDT)

Dichlorodiphenyltrichloroethane (DDT) was the first of a long line of chlorinated hydrocarbon insecticides (figure 5.4.1.n). These compounds are chains of carbon and hydrogen with chlorine atoms replacing some of the hydrogen atoms. Introduced during World War II, DDT, along with penicillin and the sulfa drugs, was responsible for the fact that this was the first war in history where trauma killed more people - combatants and noncombatants alike - than infectious disease.

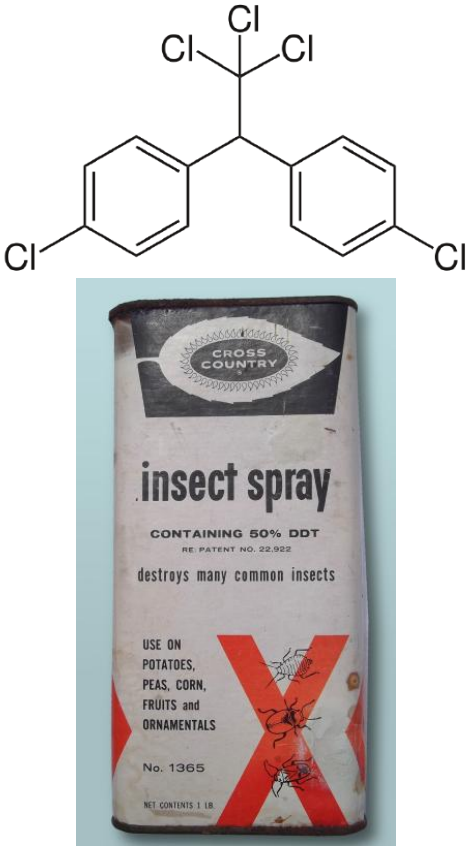


Figure 5.4.1.n: DDT was a commonly-used pesticide. The bag of DDT on the right says, "Insect spray containing 50% DDT. Destroys many common insects. Use on potatoes, peas, corn, fruits, and ornamentals." Left image by Leyo (public domain), and right image by Xanthis (public domain).

Dichlorodiphenyltrichloroethane is effective against many crop pests as well as vectors of human diseases such as the mosquitoes that spread malaria and yellow fever and fleas, which transmit the plague. Prior to the introduction of DDT, the number of cases of malaria in Ceylon (now Sri Lanka) was more than a million a year. By 1963 the disease had been practically eliminated from the island. However, growing concern about the hazards of DDT led to its abandonment there in the mid-1960s, and soon thereafter malaria became common once again. Because it remains in the environment and is resistant to breakdown, DDT was especially effective against malarial mosquitoes. One or two sprays a year on the walls of homes kept them free of mosquitoes. It was also inexpensive, further adding to its appeal, but DDT has several serious drawbacks.

Because DDT builds up in fatty tissues (bioaccumulation) and becomes more concentrated at the highest levels of the food chain (biomagnification), it is especially harmful to apex predators, such as Bald Eagles (figure 5.4.1.o). Classical studies documenting these effects were described in the 1960s bestseller *Silent Spring* by Rachel Carson. It was discovered that DDT caused the eggshells of birds to become fragile and break, making reproduction impossible. As a result, the Bald Eagle was listed as an endangered species under U.S. law. After DDT was banned in the United States in 1972, affected bird populations made noticeable recoveries, including the iconic Bald Eagle.



Figure 5.4.1.*o*: DDT pesticide residue in fish and other prey poisoned Bald Eagles, causing egg shell thinning that resulted in widespread nesting failures. Image by Ron Holmes/USFWS (public domain).

Physical Hazards

Physical hazards are additional forces that can imperil humans. Physical hazards may arise naturally such as natural disasters (earthquakes, wildfires, landslides, etc.) or extreme weather (figure 5.4.1.*p*). Others may arise from human structures or activities (traffic accident, building collapse, injury from mechanical equipment, strain on the body from repeated movements, etc.) Some physical hazards, such as explosions or radiation, can arise from natural or human sources.



Figure 5.4.1.*p*: This section of E. Grace Street, Richmond, VA, collapsed during tropical storm Gaston. Gaston dropped twelve inches of rain in the area. Image by Liz Roll/FEMA Photo Library (public domain).

Radiation is energy given off by matter in the form of rays or high-speed particles, and some types of radiation present a physical hazard. Some familiar forms of radiation are infrared radiation (heat), visible light, ultraviolet (UV) light, radio waves, and microwaves. We are exposed to radiation every day from natural sources. For example, the sun exposes us to UV radiation. We are also exposed to radiation from human-made sources like medical X-rays and smoke detectors. We're even exposed to low levels of radiation on cross-country flights, from watching television, and even from some construction materials. Some types of radioactive materials are more dangerous than others. Specifically, **ionizing radiation**, like X rays and gamma rays (one of the forms of radiation emitted from nuclear fuel and waste), have enough energy to break molecular bonds and displace (or remove) electrons from atoms.

Attribution

Modified by Melissa Ha from the following sources:

- *Environmental Biology* by Mathew R. Fisher (licensed under CC-BY)

- The Impacts of Environmental Conditions
- Environmental Health
- Environmental Toxicology
- Pests and Pesticides
- Energy Flow through Ecosystems

- Insecticides from *Biology* by John W. Kimball (licensed under CC-BY)
- Impact on Human Health from *AP Environmental Science* by University of California College Prep (licensed under CC-BY)
- Environmental Health at a glance. 2003. World Bank. Washington, DC. (licensed under CC-BY)
- Asbestos. 2018. Environment Protection Agency. Accessed 01-07-2021. (public domain)
- Per- and Polyfluoroalkyl Substances (PFAS). 2018. Environment Protection Agency. Accessed 01-07-2021. (public domain)
- Polychlorinated Biphenyls (PCBs). 2020. Environment Protection Agency. Accessed 01-08-2021. (public domain)
- What are DALYs? National Institute of Mental Health. Accessed 01-06-2021. (public domain)
- Radiation Basics. 2020. U.S. NRC. Accessed 01-07-2021 (public domain)
- Phthalates. 2017. National Biomonitoring Program. Centers for Disease Control. Accessed 01-08-2021. (public domain)

This page titled 5.4.1: Types of Environmental Hazards is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **15.1: Types of Environmental Hazards** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

5.4.2: Epidemiology

The field of **epidemiology** studies the geographical distribution and timing of infectious disease occurrences and how they are transmitted and maintained in nature, with the goal of recognizing and controlling outbreaks. The science of epidemiology includes **etiology** (the study of the causes of disease) and investigation of disease transmission (mechanisms by which a disease is spread). **Epidemiologists** are thus scientists who study the causes and patterns of human disease, which involves examining statistics to identify health threats and recommending strategies to reduce these threats.

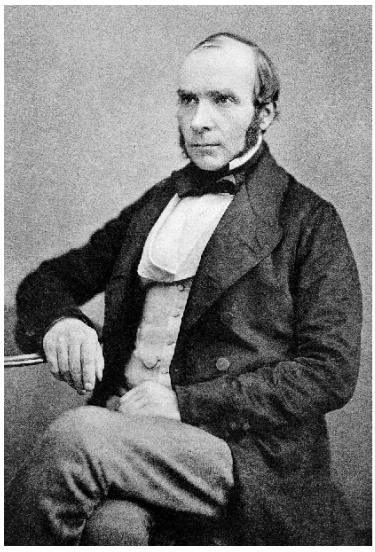
The work of epidemiologist Alice Wang is highlighted in the video below. While many of the examples on this page focus on infectious disease (which result from biological hazards), epidemiologists can study noncommunicable diseases too (such as poisoning or obesity). The text at the end of the video says, "Are You CDC? For more information bout employment opportunities with CDC, please visit: jobs.cdc.gov".



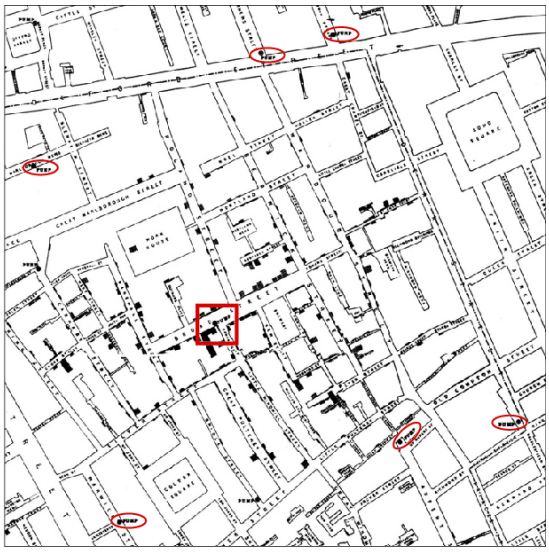
History of Epidemiology

The studies of 19th century physicians and researchers such as John Snow, Florence Nightingale, Ignaz Semmelweis, Joseph Lister, Robert Koch, Louis Pasteur, and others sowed the seeds of modern epidemiology.

John Snow (figure 5.4.2.*a*) was a British physician known as the father of epidemiology for determining the source of the 1854 Broad Street cholera epidemic in London. Based on observations he had made during an earlier cholera outbreak (1848–1849), Snow proposed that cholera was spread through a fecal-oral route of transmission and that a microbe was the infectious agent. He investigated the 1854 cholera epidemic in two ways. First, suspecting that contaminated water was the source of the epidemic, Snow identified the source of water for those infected. He found a high frequency of cholera cases among individuals who obtained their water from the River Thames downstream from London. This water contained the refuse and sewage from London and settlements upstream. He also noted that brewery workers did not contract cholera and on investigation found the owners provided the workers with beer to drink and stated that they likely did not drink water. Second, he also painstakingly mapped the incidence of cholera and found a high frequency among those individuals using a particular water pump located on Broad Street. In response to Snow's advice, local officials removed the pump's handle, resulting in the containment of the Broad Street cholera epidemic. John Snow's own account of his work has additional links and information.



(a)



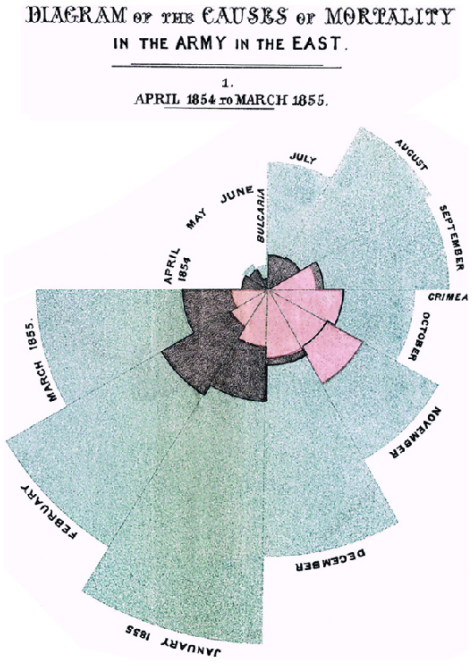
(b)

Figure 5.4.2.*a*: (a) John Snow (1813–1858), British physician and father of epidemiology. (b) Snow's detailed mapping of cholera incidence led to the discovery of the contaminated water pump on Broad street (red square) responsible for the 1854 cholera epidemic. (credit a: modification of work by "Rsabbatini"/Wikimedia Commons)

Snow's work represents an early epidemiological study and it resulted in the first known public health response to an epidemic. Snow's meticulous case-tracking methods are now common practice in studying disease outbreaks and in associating new diseases with their causes. His work further shed light on unsanitary sewage practices and the effects of waste dumping in the Thames. Additionally, his work supported the germ theory of disease, which argued disease could be transmitted through contaminated items, including water contaminated with fecal matter.

Florence Nightingale's work is another example of an early epidemiological study. In 1854, Nightingale was part of a contingent of nurses dispatched by the British military to care for wounded soldiers during the Crimean War. Nightingale kept meticulous records regarding the causes of illness and death during the war. Her recordkeeping was a fundamental task of what would later become the science of epidemiology. Her analysis of the data she collected was published in 1858. In this book, she presented monthly frequency data on causes of death in a wedge chart histogram (figure 5.4.2.*b*). This graphical presentation of data, unusual at the time, powerfully illustrated that the vast majority of casualties during the war occurred not due to wounds sustained in action but to what Nightingale deemed preventable infectious diseases. Often these diseases occurred because of poor sanitation and lack of access to hospital facilities. Nightingale's findings led to many reforms in the British military's system of medical care. Learn more about Nightingale's wedge chart here.





(a) (b)

Figure 5.4.2.*b*: (a) Florence Nightingale reported on the data she collected as a nurse in the Crimean War. (b) Nightingale's diagram shows the number of fatalities in soldiers by month of the conflict from various causes. The total number dead in a particular month is equal to the area of the wedge for that month. The colored sections of the wedge represent different causes of death: wounds (pink), preventable infectious diseases (gray), and all other causes (brown).

Joseph Lister provided early epidemiological evidence leading to good public health practices in clinics and hospitals. Most physicians did not wash their hands between patient visits or clean and sterilize their surgical tools. Lister, however, discovered the disinfecting properties of carbolic acid, also known as phenol. He introduced several disinfection protocols that dramatically lowered post-surgical infection rates. He demanded that surgeons who worked for him use a 5% carbolic acid solution to clean their surgical tools between patients, and even went so far as to spray the solution onto bandages and over the surgical site during operations (figure 5.4.2.*c*). He also took precautions not to introduce sources of infection from his skin or clothing by removing his coat, rolling up his sleeves, and washing his hands in a dilute solution of carbolic acid before and during the surgery.

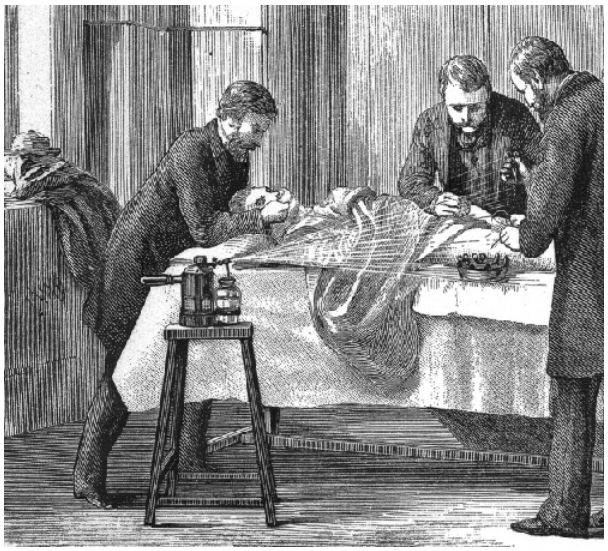


Figure 5.4.2.*c*: Joseph Lister initiated the use of a carbolic acid (phenol) during surgeries. This illustration of a surgery shows a pressurized canister of carbolic acid being sprayed over the surgical site.

Analyzing Disease in a Population

Epidemiological analyses are always carried out with reference to a population, which is the group of individuals that are at risk for the disease or condition. The population can be defined geographically, but if only a portion of the individuals in that area are susceptible, additional criteria may be required. Susceptible individuals may be defined by particular behaviors, such as intravenous drug use, owning particular pets, or membership in an institution, such as a college. Being able to define the population is important because most measures of interest in epidemiology are made with reference to the size of the population.

The state of being diseased is called **morbidity**. Morbidity in a population can be expressed in a few different ways. Morbidity, or **total morbidity,** is expressed in numbers of individuals without reference to the size of the population. The **morbidity rate** can be expressed as the number of diseased individuals out of a standard number of individuals in the population, such as 100,000, or as a percent of the population.

There are two aspects of morbidity that are relevant to an epidemiologist: a disease's prevalence and its incidence. **Prevalence** is the number, or proportion, of individuals with a particular illness in a given population at a point in time. For example, the Centers for Disease Control and Prevention (CDC) estimated that about 1.2 million people in the United States lived with an active human immunodeficiency virus (HIV) infection in 2018. Expressed as a proportion, or rate, this is a prevalence of 367 infected persons per 100,000 in the population. On the other hand, **incidence** is the number or proportion of new cases in a period of time. For the same year, the CDC estimates that there were 36,400 newly diagnosed cases of HIV infection, which is an incidence of 11.1 new cases per 100,000 in the population. The relationship between incidence and prevalence can be seen in figure 5.4.2.*d*. For a chronic disease like HIV infection, prevalence will generally be higher than incidence because it represents the cumulative number of new cases over many years minus the number of cases that are no longer active (because the patient died or was cured).

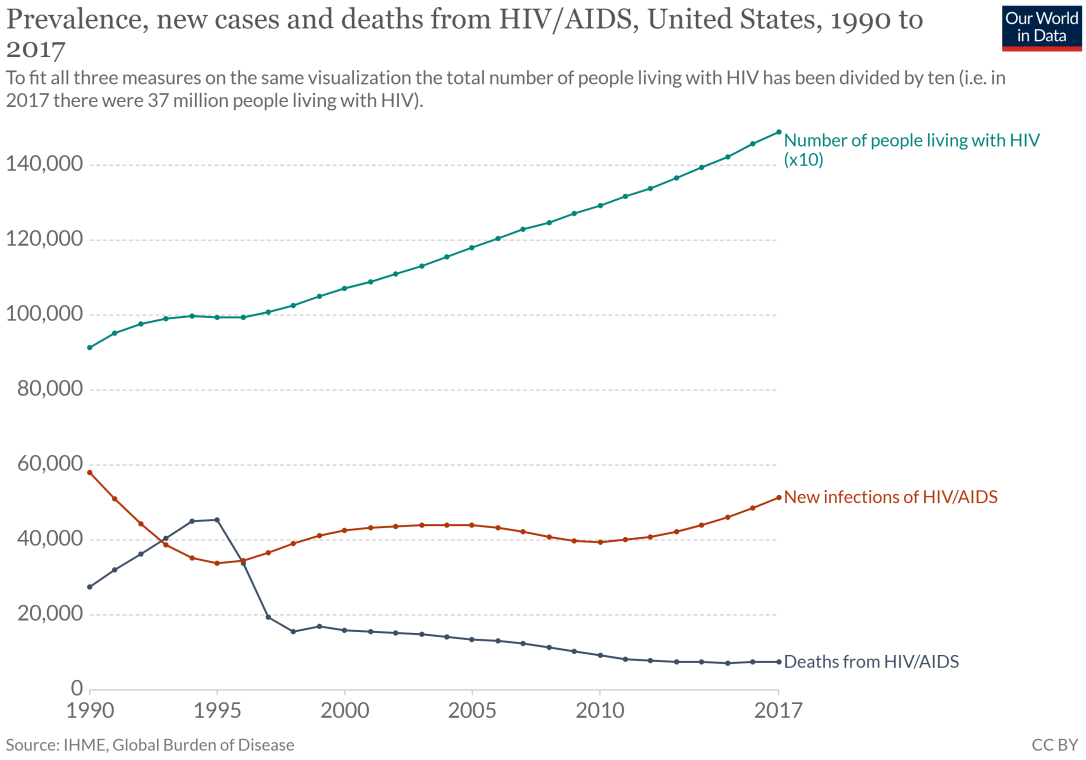


Figure 5.4.2.*d*: The number of people living with HIV (prevalence), new infections of HIV/AIDS (incidence), and deaths from HIV/AIDs in the United States from 1990 to 2017. Prevalence and incidence can also be expressed as a rate or proportion for a given population. The number of people living with HIV has increased over time, but the number of new infections has fluctuated. Deaths reached a high around 1995 and have since declined. Image by Max Roser and Hannah Ritchie (2018) - "HIV / AIDS". Published online at OurWorldInData.org. (CC-BY).

In addition to morbidity rates, the incidence and prevalence of **mortality** (death) may also be reported. A mortality rate can be expressed as the percentage of the population that has died from a disease or as the number of deaths per 100,000 persons (or other suitable standard number).

Diseases that are seen only occasionally, and usually without geographic concentration, are called **sporadic diseases**. Examples of sporadic diseases include tetanus, rabies, and plague. In the United States, *Clostridium tetani*, the bacterium that causes tetanus, is ubiquitous in the soil environment, but incidences of infection occur only rarely and in scattered locations because most individuals are vaccinated, clean wounds appropriately, or are only rarely in a situation that would cause infection. Likewise in the United States there are a few scattered cases of plague each year, usually contracted from rodents in rural areas in the western states.

Diseases that are constantly present (often at a low level) in a population within a particular geographic region are called **endemic diseases**. For example, malaria is endemic to some regions of Brazil, but is not endemic to the United States.

Diseases for which a larger than expected number of cases occurs in a short time within a geographic region are called **epidemic diseases**. Influenza is a good example of a commonly epidemic disease. Incidence patterns of influenza tend to rise each winter in the northern hemisphere. These seasonal increases are expected, so it would not be accurate to say that influenza is epidemic every winter; however, some winters have an usually large number of seasonal influenza cases in particular regions, and such situations would qualify as epidemics (figures 5.4.2.$e-f$).

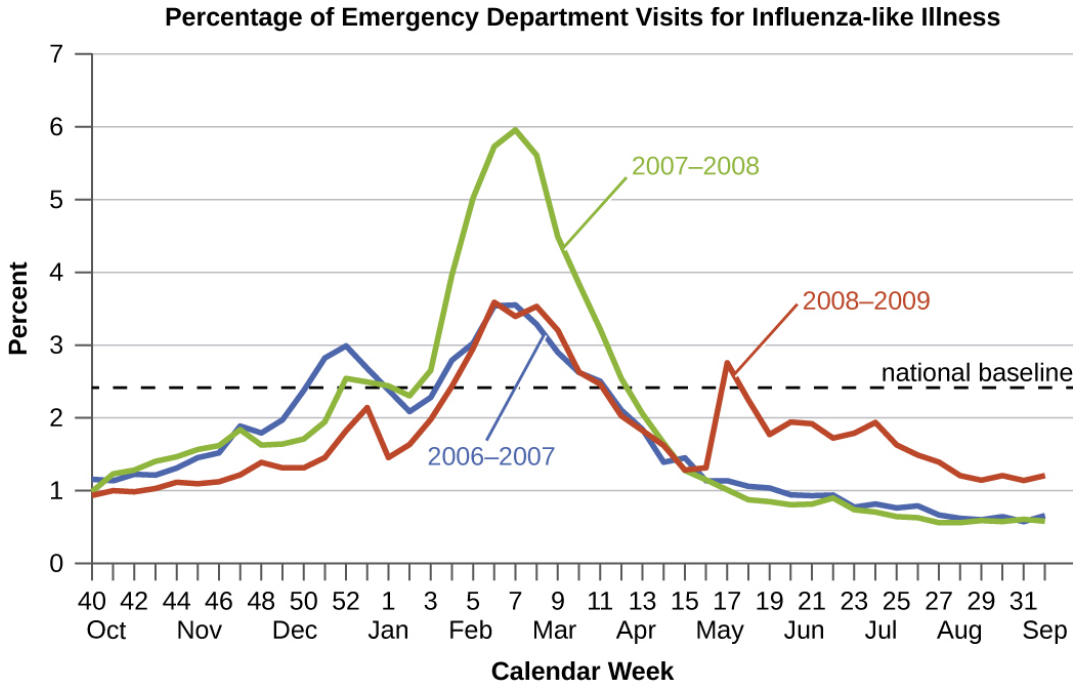


Figure 5.4.2.e: The 2007–2008 influenza season in the United States saw much higher than normal numbers of visits to emergency departments for influenza-like symptoms as compared to the previous and the following years. The national baseline is near 2.5%. All years have a small peak in January and a larger peak February to April. 2007-2008 had the largest peak in February to April. 2008 – 2009 had an additional peak from May to September. (credit: modification of work by Centers for Disease Control and Prevention)

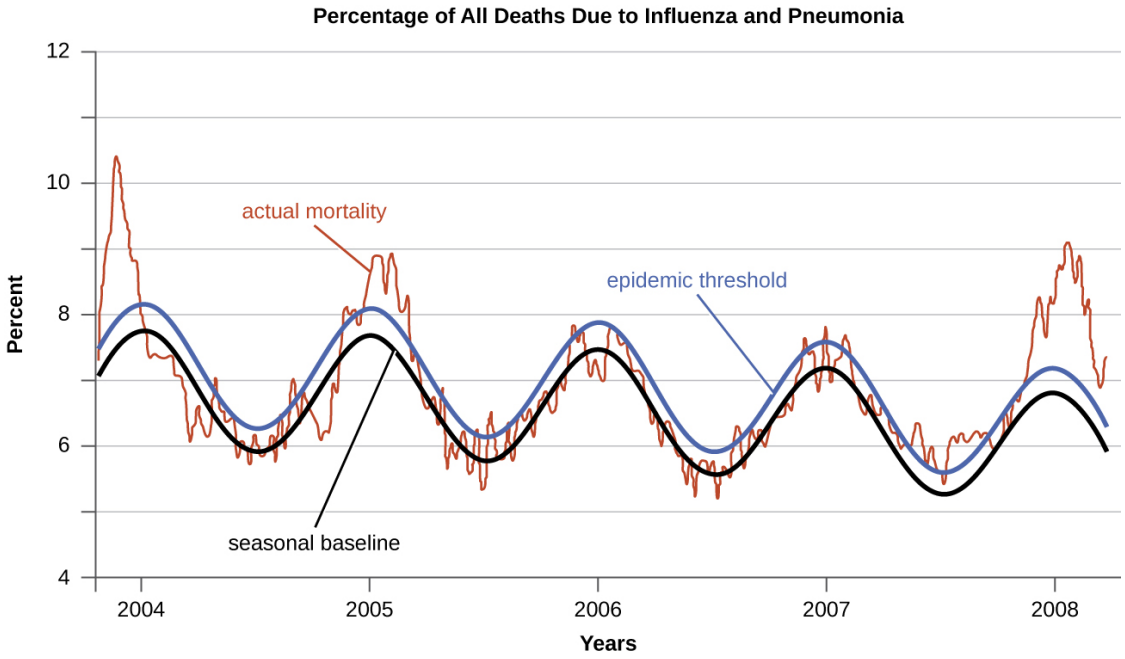


Figure 5.4.2.*f*: The seasonal epidemic threshold (blue curve) is set by the CDC-based data from the previous five years. The seasonal baseline fluctuates to a high in winter. The epidemic threshold is just a half percent higher than the baseline. The actual mortality fluctuates above and below both of these lines. When actual mortality rates exceed this threshold, a disease is considered to be epidemic. As this graph shows, pneumonia- and influenza-related mortality saw pronounced epidemics during the winters of 2003–2004, 2005, and 2008. (credit: modification of work by Centers for Disease Control and Prevention)

An epidemic that occurs on a worldwide scale is called a **pandemic disease**. For example, HIV/AIDS (acquired immunodeficiency syndrome) and coronavirus disease 19 (COVID-19) are pandemic diseases, and novel influenza virus strains often become pandemic.

Etiology

When studying an epidemic, an epidemiologist's first task is to determinate the cause of the disease, called the etiologic agent or causative agent. Connecting a disease to a specific pathogen can be challenging because of the extra effort typically required to demonstrate direct causation as opposed to a simple association. It is not enough to observe an association between a disease and a suspected pathogen; controlled experiments are needed to eliminate other possible causes. In addition, pathogens are typically difficult to detect when there is no immediate clue as to what is causing the outbreak. Signs and symptoms of disease are also commonly nonspecific, meaning that many different agents can give rise to the same set of signs and symptoms. This complicates diagnosis even when a causative agent is familiar to scientists.

Robert Koch was the first scientist to specifically demonstrate the causative agent of a disease (anthrax) in the late 1800s. Koch developed four criteria, now known as Koch's postulates, which had to be met in order to positively link a disease with a pathogenic microbe (figure 5.4.2.*g*). Between 1876 and 1905, many common diseases were linked with their etiologic agents, including cholera, diphtheria, gonorrhea, meningitis, plague, syphilis, tetanus, and tuberculosis. Today, we use the molecular Koch's postulates, a variation of Koch's original postulates. (You can read more about both here.)

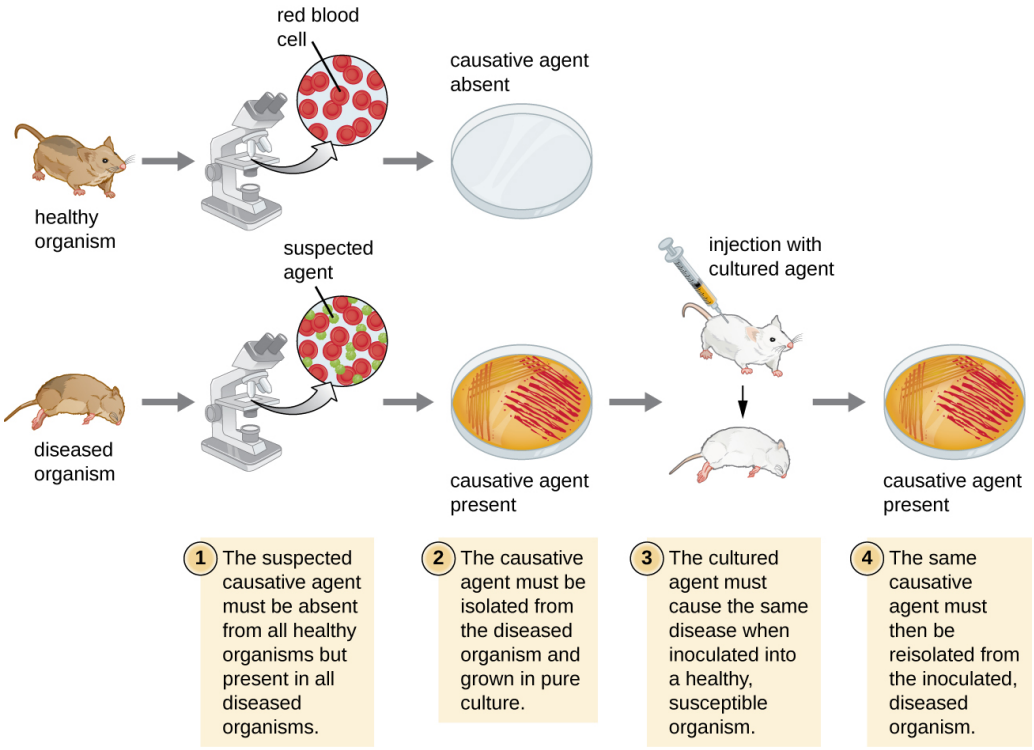


Figure 5.4.2.*g*: The steps for confirming that a pathogen is the cause of a particular disease using Koch's postulates. (1) Only diseased organisms should contain the suspected causative agent (potential pathogen). (2) The causative agent from diseased organisms must be cultured (grown in the lab). (3) The cultured agent must then cause the same disease when a healthy, susceptible individual is infected with it (inoculated). (4) The same causative agent should be isolated and cultured from the individual in the previous step. Image and caption (modified) from OpenStax (CC-BY).

How Diseases Spread

Understanding how infectious pathogens spread is critical to preventing infectious disease. For pathogens to persist over long periods of time they require **reservoirs** where they normally reside. Reservoirs can be living organisms (rats, bats, racoons, livestock, etc.) or nonliving sites. Nonliving reservoirs can include soil and water in the environment. These may naturally harbor the organism because it may grow in that environment. These environments may also become contaminated with pathogens in human feces, pathogens shed by intermediate hosts, or pathogens contained in the remains of intermediate hosts.

An individual capable of transmitting a pathogen without displaying symptoms is referred to as a **carrier**. A **passive carrier** is contaminated with the pathogen and can mechanically transmit it to another host; however, a passive carrier is not infected. By contrast, an **active carrier** is an infected individual who can transmit the disease to others. An active carrier may or may not exhibit signs or symptoms of infection.

Regardless of the reservoir, transmission must occur for an infection to spread. **Contact transmission** includes **direct contact** or **indirect contact**. Person-to-person transmission is a form of direct contact transmission. Here the agent is transmitted by physical contact between two individuals (figure 5.4.2.*h*) through actions such as touching, kissing, sexual intercourse, or droplet sprays. Indirect contact transmission involves inanimate objects called **fomites** that become contaminated by pathogens from an infected individual or reservoir (figure 5.4.2.*i*). For example, an individual with the common cold may sneeze, causing droplets to land on a fomite such as a tablecloth or carpet, or the individual may wipe her nose and then transfer mucus to a fomite such as a doorknob or towel.




Figure 5.4.2.*h*: Direct contact transmission of pathogens can occur through physical contact. Many pathogens require contact with a mucous membrane to enter the body, but the host may transfer the pathogen from another point of contact (for example, hand) to a mucous membrane (for example, mouth or eye). (credit left: modification of work by Lisa Doehnert)

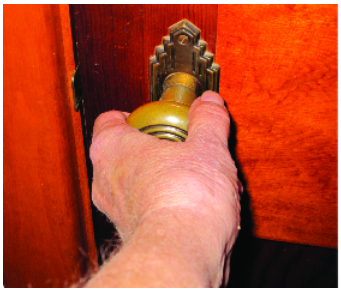





Figure 5.4.2.*i*: Fomites are nonliving objects that facilitate the indirect transmission of pathogens. Contaminated doorknobs, towels, and syringes are all common examples of fomites. (credit left: modification of work by Kate Ter Haar; credit middle: modification of work by Vernon Swanepoel; credit right: modification of work by "Zaldylmg"/Flickr)

Vehicle transmission refers to the transmission of pathogens through vehicles such as water, food, and air. Water contamination through poor sanitation methods leads to waterborne transmission of disease. **Waterborne** disease remains a serious problem in many regions throughout the world. The World Health Organization (WHO) estimates that contaminated drinking water is responsible for more than 500,000 deaths each year. Similarly, food contaminated through poor handling or storage can lead to **foodborne** transmission of disease. Dust and fine particles known as **aerosols**, which can float in the air, can carry pathogens and facilitate the **airborne** transmission of disease.

Diseases can also be transmitted by a mechanical or biological **vector**, an animal (typically an arthropod) that carries the disease from one host to another. Ticks, fleas, and mosquitos are examples of vectors (figure 5.4.2.*j*). **Mechanical transmission** is facilitated by a **mechanical vector**, an animal that carries a pathogen from one host to another without being infected itself. **Biological transmission** occurs when the pathogen reproduces within a **biological vector** that transmits the pathogen from one host to another (figure 5.4.2.*k*).

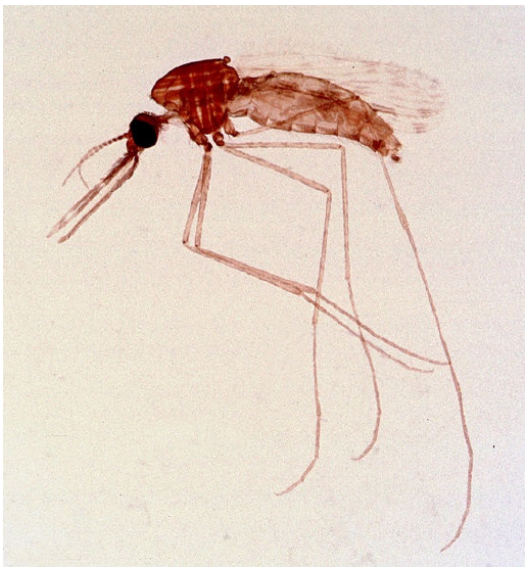


Figure 5.4.2.j: The *Anopheles* mosquito is the **vector** for *Plasmodium*, which causes malaria. Image by Alan R Walker (CC-BY-SA).

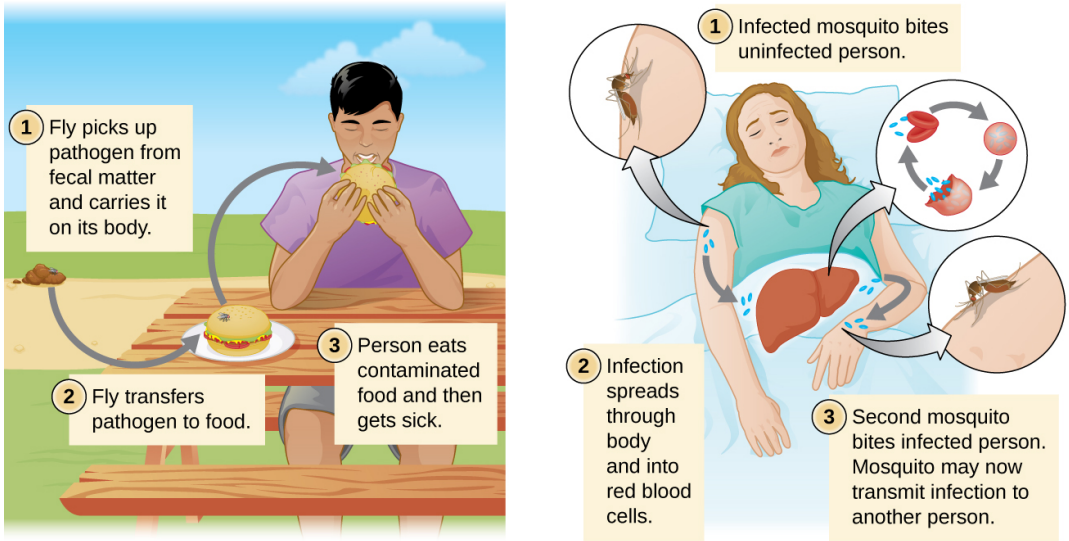


Figure 5.4.2.k: (a) A mechanical vector carries a pathogen on its body from one host to another, not as an infection. Step 1: fly picks up pathogen from fecal matter and carries it on its body. 2: Fly transfers pathogen to food. 3: Person eats contaminated food and gets sick. (b) A biological vector carries a pathogen from one host to another after becoming infected itself. Step 1: Infected mosquito bites uninfected person. 2: Infections spreads through body and into red blood cells. 3: Second mosquito bites infected person. Mosquito may now transmit infection to another person.

Ecosystem degradation can increase disease spread through several mechanisms. First, it can promote vector populations. For example, deforestation, dams, and urbanization increase the amount of standing water, increasing habitat for vectors, such as mosquitoes. Additionally, climate change may expand the range of disease vectors that are limited to tropical or subtropical regions. Second, disruptions in the water cycle can create conditions that favor pathogens. For example, fertilizer runoff from agriculture increases nutrient levels, making bodies of water more suitable for the bacterium that causes cholera. Increased flooding as a result of changes in land use and climate change also facilitates the spread of waterborne disease.

A third way that ecosystem degradation can increase disease spread is through decreasing biodiversity. For example, Lyme disease is caused by the bacterium *Borrelia* and is transmitted to humans from small mammals (the reservoirs) by ticks (figures 5.4.2.$l-m$). Some reservoirs (squirrels and opossums) remove ticks, reducing infection, but field mice do not. Habitat fragmentation has increased field mice relative to squirrels and opossums and thus increased the reservoir for Lyme disease.

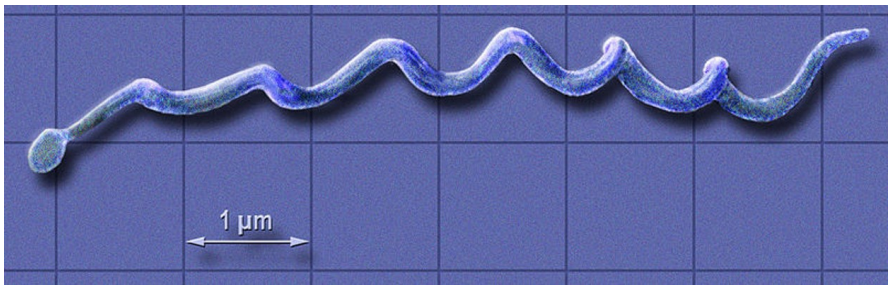


Figure 5.4.2.*l*: Drawing of the bacterium *Borrelia burgdorferi*, which causes Lyme disease. Image by Lamiot (CC-BY-SA).

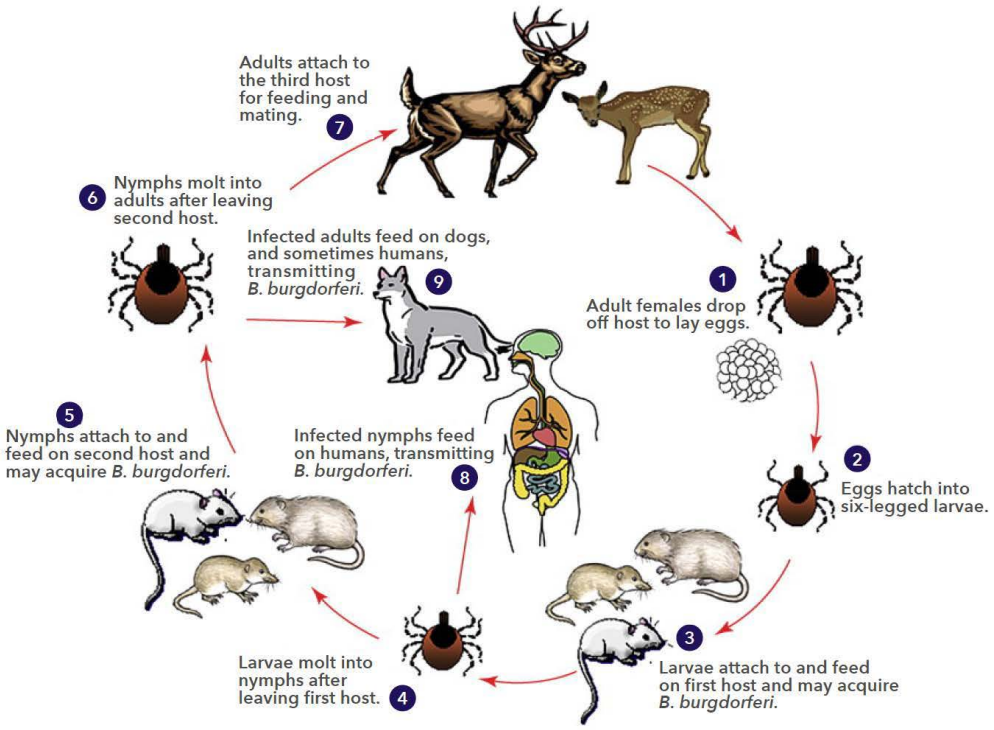


Figure 5.4.2.*m*: Life cycle of the bacterium *Borrelia burgdorferi*, which causes Lyme disease. Ticks can acquire this bacterium when the feed on the initial host and then transmit the disease to humans. Step 1: Adult females drop off host to lay eggs. 2: Eggs hatch into six-legged larvae. 3: Larvae attach to and feed on first host (various small mammals) and may acquire *B. burgdorferi*. 4: Larvae molt into nymphs after leaving first host. 5: Nymphs attach to and feed on second host and may acquire *B. burgdorferi*. 6: Nymphs molt into adults after leaving second host. 7: Adults attach to the third host (such as deer) for feeding and mating. 8: An arrow branches from step 4 indicating that infected nymphs feed on humans, transmitting *B. burgdorferi*. 9: Infected adults feed on dogs and sometimes humans, transmitting *B. burgdorferi*. Image from Tick-Borne Disease Working Group 2018 Report to Congress.

Individuals suspected or known to have been exposed to certain contagious pathogens may be **quarantined**, or isolated to prevent transmission of the disease to others. Hospitals and other health-care facilities generally set up special wards to isolate patients with particularly hazardous diseases such as tuberculosis or COVID-19 (figure 5.4.2.*n*). Depending on the setting, these wards may be equipped with special air-handling methods, and personnel may implement special protocols to limit the risk of transmission, such as personal protective equipment or the use of chemical disinfectant sprays upon entry and exit of medical personnel.



(a) (b)

Figure 5.4.2.n: (a) The Aeromedical Biological Containment System (ABCS) is a module designed by the CDC and Department of Defense specifically for transporting highly contagious patients by air. (b) An isolation ward for Ebola patients in Lagos, Nigeria. (credit a: modification of work by Centers for Disease Control and Prevention; credit b: modification of work by CDC Global)

Attribution

Modified by Melissa Ha from Disease and Epidemiology from *Microbiology* by OpenStax (licensed under CC-BY). Access for free at openstax.org.

This page titled 5.4.2: Epidemiology is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **15.2: Epidemiology** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

5.4.3: Infectious Diseases

Worldwide, infectious diseases (biological hazards) accounted for three of the 10 leading causes of death in 2020. In children under five leading causes of death include acute respiratory infections (from indoor air pollution); diarrheal diseases (mostly from poor water, sanitation, and hygiene); and other infectious diseases such as malaria. Children are especially susceptible to environmental factors that put them at risk of developing illness early in life. **Malnutrition** (the condition that occurs when body does not get enough nutrients) is an important contributor to child mortality—malnutrition and environmental infections are inextricably linked.

The World Health Organization (WHO) recently concluded that about 50% of the consequences of malnutrition are in fact caused by inadequate water and sanitation provision and poor hygienic practices. For example, intestinal worms, which thrive in poor sanitary conditions, infect close to 90 percent of children in the developing world and, depending on the severity of the infection may lead to malnutrition, anemia, or stunted growth. About 6 million people are blind from trachoma, a disease caused by the lack of clean water combined with poor hygiene practices.

An **emerging infectious disease** is either new to the human population or has shown an increase in prevalence in the previous twenty years. Examples include coronavirus disease 19 (COVID-19), Ebola virus disease, West Nile virus, Zika virus, sudden acute respiratory syndrome (SARS), H1N1 influenza; swine and avian influenza (swine, bird flu), human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDs), and a variety of other viral, bacterial, and protozoal diseases (table 5.4.3.*a*). Whether the disease is new or conditions have changed to cause an increase in frequency, its status as emerging implies the need to apply resources to understand and control its growing impact. Emerging diseases may change their frequency gradually over time, or they may experience sudden widespread growth.

Table 5.4.3.*a***:** Some Emerging and Reemerging Infectious Diseases

Disease	Pathogen	Year Discovered	Affected Regions	Transmission
AIDS	HIV	1981	Worldwide	Contact with infected body fluids
Chikungunya fever	Chikungunya virus	1952	Africa, Asia, India; spreading to Europe and the Americas	Mosquito-borne
Ebola virus disease	Ebola virus	1976	Central and Western Africa	Contact with infected body fluids
H1N1 Influenza (swine flu)	H1N1 virus	2009	Worldwide	Droplet transmission
Lyme disease	*Borrelia burgdorferi* bacterium	1981	Northern hemisphere	From mammal reservoirs to humans by tick vectors
West Nile virus disease	West Nile virus	1937	Africa, Australia, Canada to Venezuela, Europe, Middle East, Western Asia	Mosquito-borne

Emerging diseases are found in all countries, both developed and developing. Some nations are better equipped to deal with them. National and international public health agencies watch for epidemics in developing countries because those countries rarely have the health-care infrastructure and expertise to deal with large outbreaks effectively. In addition to the altruistic goal of saving lives and assisting nations lacking in resources, the global nature of transportation means that an outbreak anywhere can spread quickly to every corner of the planet. Managing an epidemic in one location—its source—is far easier than fighting it on many fronts.

In 2015, the WHO set priorities on several emerging diseases that had a high probability of causing epidemics and that were poorly understood (and thus urgently required research and development efforts).

A **reemerging infectious disease** is a disease that is increasing in frequency after a previous period of decline. Its reemergence may be a result of changing conditions or old prevention regimes that are no longer working. Examples of such diseases are drug-

resistant forms of tuberculosis, bacterial pneumonia, and malaria. Drug-resistant strains of the bacteria causing gonorrhea and syphilis are also becoming more widespread, raising concerns of untreatable infections.

A variety of environmental factors may contribute to re-emergence of a particular disease, including temperature, moisture, human food or animal feed sources, etc. Disease re-emergence may be caused by the coincidence of several of these environmental and/or social factors to allow optimal conditions for transmission of the disease.

It seems likely that a wide variety of infectious diseases have affected human populations for thousands of years emerging when the environmental, host, and agent conditions were favorable. Expanding human populations have increased the potential for transmission of infectious disease as a result of close human proximity and increased likelihood for humans to be in “the wrong place at the right time” for disease to occur (eg, natural disasters or political conflicts). Global travel increases the potential for a carrier of disease to transmit infection thousands of miles away in just a few hours, as evidenced by WHO precautions concerning international travel and health.

Selected Infectious Diseases

COVID-19

Coronavirus disease 19 (**COVID-19**) is a respiratory illness caused by a virus called novel species of coronavirus (SARS-CoV-2). Symptoms include fever or chills, cough, shortness or breath or difficulty breathing, fatigue, headache, and sore throat. The WHO declared coronavirus disease 19 (COVID-19) a public health emergency of international concern on January 30 and a pandemic on March 11, 2020. A seemingly limited cluster of cases of pneumonia linked to a sea food market in Wuhan, China has become one of the worst pandemics in human history with a staggering number of more than 90.2 million infections in 177 countries and 1.9 million deaths globally as of January 10, 2021 (figure 5.4.3.*a*). This illustrates how locally emerging pathogens have the capacity to spread rapidly and cross borders and become an imminent public health threat to the entire world. You can view the the latest COVID-19 data here.

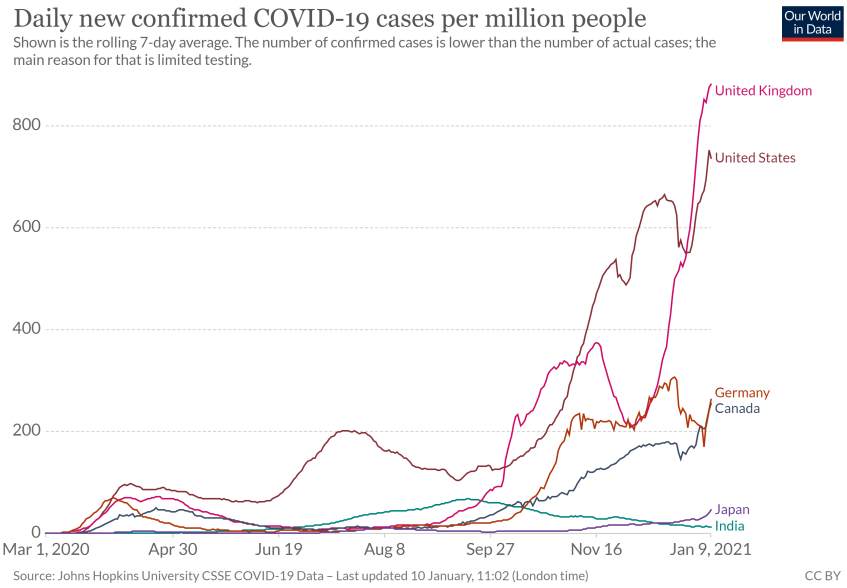


Figure 5.4.3.*a*: New confirmed COVID-19 cases per million people from March 1, 2020 to Junuary 10, 2021 in selected countries. Image by Max Roser, Hannah Ritchie, Esteban Ortiz-Ospina and Joe Hasell (2020) - "Coronavirus Pandemic (COVID-19)". Published online at OurWorldInData.org. (CC-BY).

Studies indicate that SARS-CoV-2 might originate from bat coronaviruses and infects humans directly or through other animals as intermediaries. It primarily spreads between humans through the respiratory tract by involving droplets, aerosols, respiratory secretions, or saliva. It is most often spread by direct contact transmission, but it airborne transmission has been documented in certain situations, such as in an enclosed space (figure 5.4.3.*b*). Researchers continue to investigate airborne transmission of the virus. This coronavirus is infectious during its incubation period, which is reported to be 3-7 days, at most 14 days, when no

symptoms are shown in patients. Those infected may experience mild to moderate disease (80%), severe disease (15%), and critical illness (5%) with an overall case fatality rate of 0.5–2.8% with much higher rates (3.7–14.8%) in octogenarians. The severe and critical illness categories (about 20% of all infections) have overwhelmed health systems worldwide (figure 5.4.3.*c*). The Centers for Disease Control and Prevention's COVID-19 page has the latest information about preventing transmission, vaccines, symptoms, and testing.

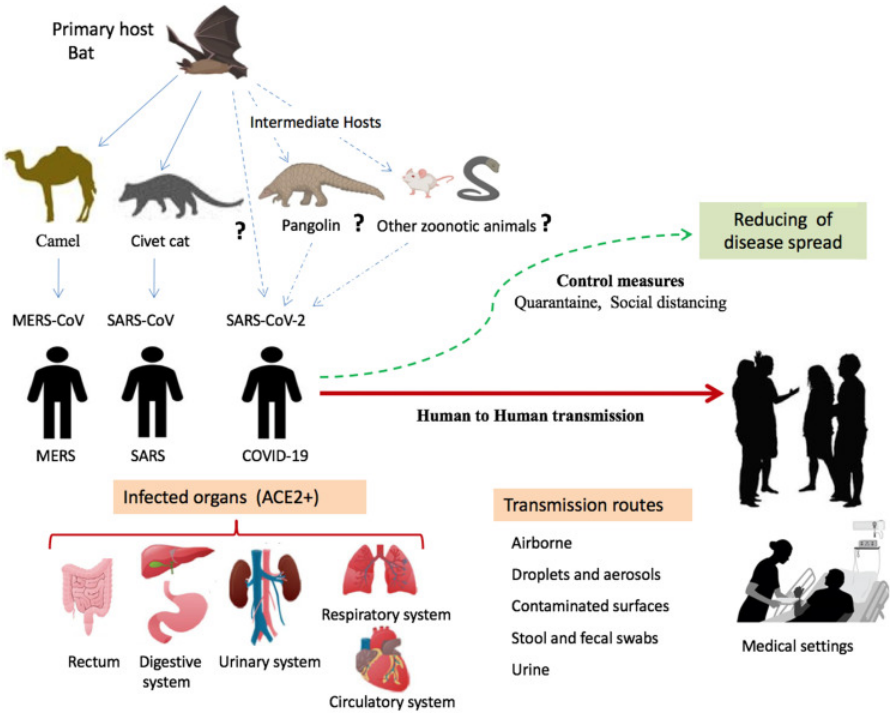


Figure 5.4.3.*b*: Origin of SARS-CoV-2 and potential transmission routes to humans. SARS-CoV-2 transmitted to humans from bats directly or by intermediate hosts such as rats, pangolins, snakes, and rats. The virus is spread among humans by different routes including droplets, aerosols, direct contact, and other potential routes such as urine, stool, and fecal swabs. Human organs with a protein called the ACE2 receptor (ACE2+) are targets for viral infection. Disease spread can be limited through quarantining when sick, social distancing, avoiding large gatherings (not shown), wearing masks (not shown), and vaccinations (not shown). Image and caption (modified) from Tizaoui, K., Zidi, I., Lee, K. H., Ghayda, R. A., Hong, S. H., Li, H., Smith, L., Koyanagi, A., Jacob, L., Kronbichler, A., & Shin, J. I. (2020). Update of the current knowledge on genetics, evolution, immunopathogenesis, and transmission for coronavirus disease 19 (COVID-19). *International journal of biological sciences*, *16*(15), 2906–2923. https://doi.org/10.7150/ijbs.48812 (CC-BY).



Figure 5.4.3.*c*: Lieutenant (junior grade) Natasha McClinton, a surgical nurse, prepares a patient for a procedure in the intensive care unit aboard the U.S. hospital ship USNS Comfort. The ship cares for critical and non-critical patients without regard to their COVID-19 status. In April of 2020, Comfort worked with Javits New York Medical Station as an integrated system to relieve the New York City medical system as a response to the COVID-19 pandemic. Image and caption (modified) from US Navy Mass Communication Specialist 2nd Class Sara Eshleman (public domain).

Ebola Virus Disease

Ebola virus disease, previously known as Ebola hemorrhagic fever, is a rare and deadly disease caused by infection with one of the Ebola virus strains (figure 5.4.3.*d*). Ebola can cause disease in humans and nonhuman primates. Health experts had been aware of the Ebola virus since the 1970s, but outbreak at scale that had not yet been seen occurred in western Africa through 2014–2016. Previous human epidemics had been small, isolated, and contained. Indeed, the gorilla and chimpanzee populations of western Africa had suffered far worse from Ebola than the human population. The pattern of small isolated human epidemics changed in 2014. Its high transmission rate, coupled with cultural practices for treatment of the dead and perhaps its emergence in an urban setting, caused the disease to spread rapidly, and thousands of people died. The international public health community responded with a large emergency effort to treat patients and contain the epidemic.

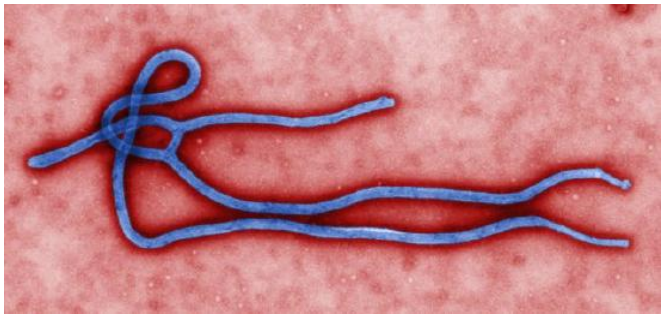


Figure 5.4.3.*d*: The Ebola virus. Image by CDC (public domain).

The 2014-2016 Ebola epidemic was the largest in history (with 28,616 cases and 11,310 deaths), affecting multiple countries in West Africa. There were a small number of cases reported in Nigeria and Mali and a single case reported in Senegal; however, these cases were contained, with no further spread in these countries. Even with the support of international agencies, the systems in western Africa struggled to identify and care for the sick and control spread. Engaging local leaders in prevention programs and messaging, along with careful policy implementation at the national and global level, helped to eventually contain the spread of the virus and put an end to this outbreak. The U.S. Food and Drug Administration approved a vaccine for Ebola late in 2019.

HIV/AIDS

Human immunodeficiency virus (**HIV**) is a virus that attacks the body's immune system. If HIV is not treated, it can lead to **acquired immunodeficiency syndrome** (**AIDS**). Those with AIDS have weakened immune systems and are thus more susceptible

to cancers and other infections. First identified in 1981, the **HIV/AIDS** epidemic has spread with ferocious speed. In 2019, 38 million people were living HIV globally. Over 95 percent of people living with HIV are in low- and middle- income countries. There were 1.7 million new infections in 2019, and more than half of new infections are among young people below age 25. In 2019 alone, 690,000 million people died from AIDS and related illnesses. In 2016, HIV/AIDS is was the second leading cause of death in Africa. At its peak, the epidemic cut life expectancy by more than 10 years in several nations (figure 5.4.3.*e*).

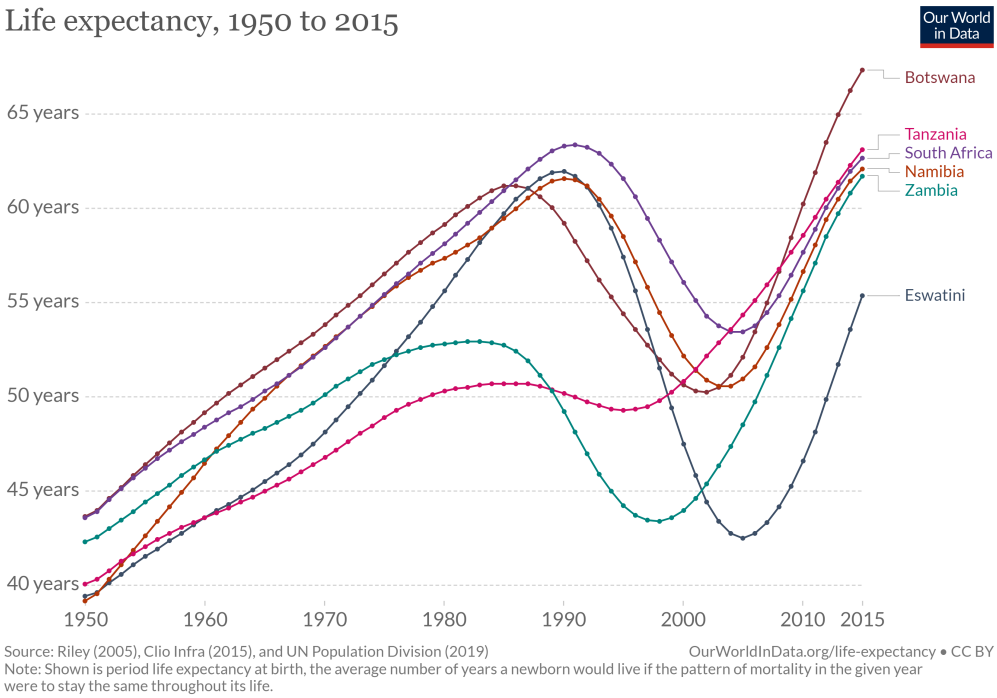


Figure 5.4.3.*e*: Life expectancy in six countries in Sub-Saharan Africa dropped dramatically in the late 1990s and early 2000s. As the efforts to control the pandemic reduced the number of new infections (incidence) and death rates from HIV, life expectancy has again increased. Image by Max Roser and Hannah Ritchie (2018) - "HIV / AIDS". Published online at OurWorldInData.org. (CC-BY).

Malaria

Approximately 40% of the world's people—mostly those living in the world's poorest countries—are at risk from malaria. **Malaria** is an infectious disease spread by mosquitoes but caused by a single-celled protozoan parasite called *Plasmodium*. There were 229 million cases of malaria and 409,000 deaths in 2019 with most cases and deaths found in Sub-Saharan Africa (**figure** 5.4.3.*f*)**.** However, Asia, Latin America, the Middle East, and parts of Europe are also affected. Pregnant women are especially at high risk of malaria. Non-immune pregnant women risk both acute and severe clinical disease, resulting in fetal loss in up to 60% of such women and maternal deaths in more than 10%, including a 50% mortality rate for those with severe disease. Semi-immune pregnant women with malaria infection risk severe anemia (lack of sufficient red blood cells to deliver oxygen) and impaired fetal growth, even if they show no signs of acute clinical disease. An estimated 10,000 women and 200,000 infants die annually as a result of malaria infection during pregnancy.

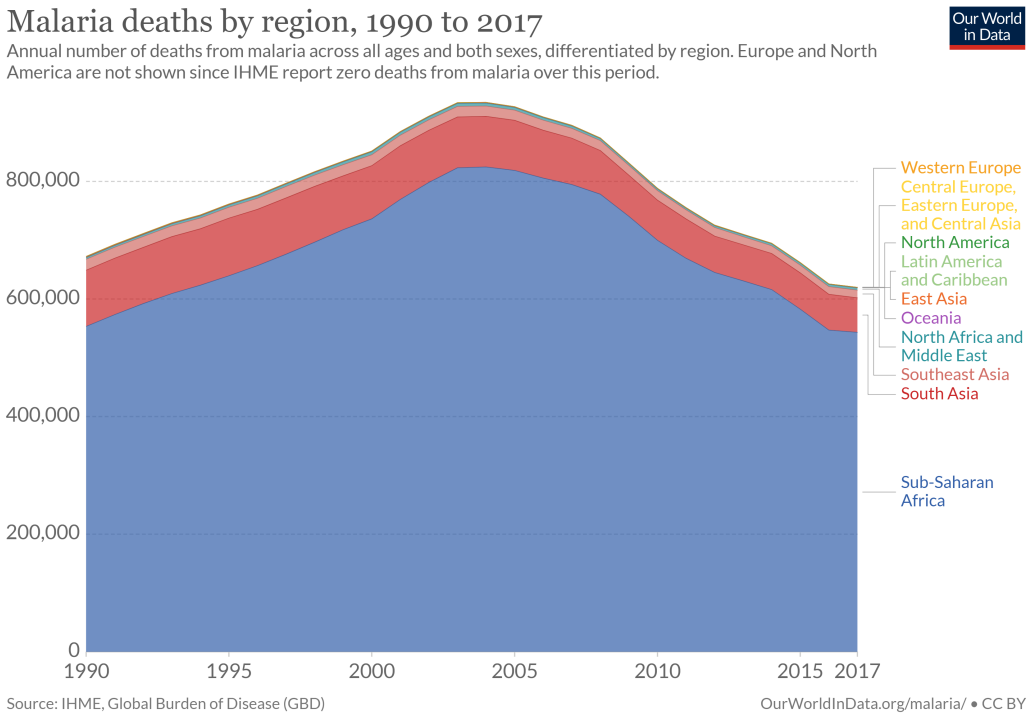


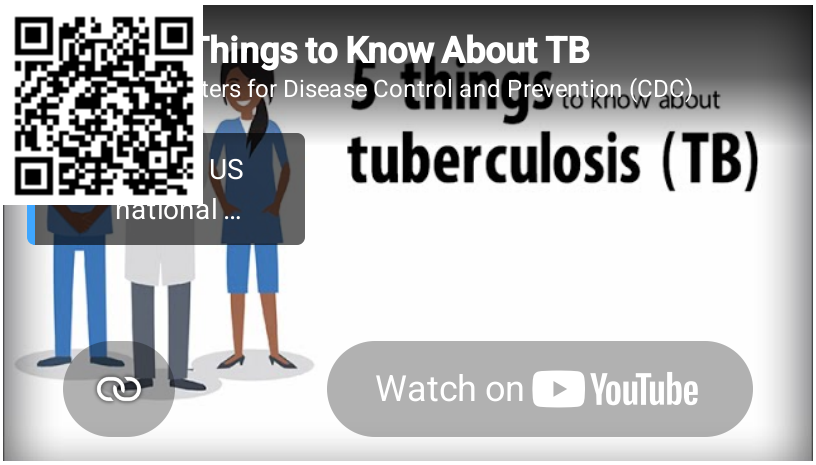
Figure 5.4.3.f: Deaths caused by malaria from 1990 to 2017. Deaths peaked above 800,000 around 2004 and were just above 600,000 in 2017. The largest share of deaths occurred in Sub-Saharan Africa. Image by Max Roser and Hannah Ritchie (2013) - "Malaria". Published online at OurWorldInData.org. (CC-BY).

Tuberculosis

The WHO estimates that 10.4 million new cases and 1.5 million deaths occur from **tuberculosis** (**TB**) each year. One-third of TB cases remain unknown to the health care system. For those accessing treatment, however, prevalence and mortality have declined significantly, and millions of lives have been saved.

Tuberculosis is caused by the bacterium *Mycobacterium tuberculosis*, which is transmitted between humans through the respiratory route and most commonly affects the lungs but can damage any tissue. When an infected individual coughs, exhales, speaks, etc., the bacterium is released in droplets that can remain suspended in the air for several hours. Only a minority (approximately 10 percent) of individuals infected with *M. tuberculosis* progress to active TB disease, while the remainder may maintain a latent infection that serves as a reservoir. TB has special challenges, including (a) a substantial number of patients with active disease are asymptomatic, capable of transmitting infection without knowing it; (b) patients must maintain compliance with treatment for six to nine months; and (c) the pathogen persists in many infected individuals in a latent state for many years but can be reactivated over a lifetime to cause disease and become transmissible.

This video by the CDC explains the basics of TB.



Disease Eradication

Many diseases have been eradicated or are nearly eradicated worldwide (Table 5.4.3.*b*). Strategies to eradicate disease infectious diseases include vaccination, improved hygiene and sanitation, preventative medication, and public health education. Some of these strategies are discussed in Environmental Health Reduction, and vaccination is discussed below.

Table 5.4.3.*b*: Infectious diseases that have been eradicated and could be eradicated in the future.

Disease	Burden of disease	Cause	Ways to eradicate	Fatality
Smallpox	Declared eradicated in 1980	Variola virus	Eradicated using vaccination	30%
Rinderpest	Declared eradicated in 2011	Rinderpest virus	Sanitary measures and vaccination	100%
Poliomielitis	116 cases in 2017	Poliovirus	Vaccination	For paralytic polio 2-5% in children and increases to 15-30% in adults
Guinea worm disease	30 cases in 2017	Parasitic worm *Dracunculus medinensis*	Hygiene, water decontamination and health education	Not fatal but debilitating
Measles	173,457 reported cases to WHO in 2017	Measles morbillivirus	Vaccination	15%
Mumps	560,622 reported cases to WHO in 2017	Mumps orthorubulavirus	Vaccination	0.01% for mumps-caused encephalitis
Rubella	6,789 reported cases to WHO in 2017	Rubella virus	Vaccination	Not reported
Lymphatic filariasis	No estimate available. In 2014, 68 million people were infected and 790 million people where at risk of infection	Roundworms: *W. bancrofti*, *B. malayi*, *B. timori*	Preventive chemotherapy	Not fatal but highly debilitating
Cysticercosis	2.56–8.30 million cases estimated by the WHO	Tapeworms: *T. solium*, *T. saginata*, *T. asiatica*	Sanitation and health education. Vaccination of pigs	Varies between countries <1-30%

Table modified from Max Roser, Sophie Ochmann, Hannah Behrens, Hannah Ritchie and Bernadeta Dadonaite (2014) - "Eradication of Diseases". Published online at OurWorldInData.org. (CC-BY).

Vaccination

Vaccines function in inducing a specific immune response that can immediately counter a pathogen if exposed. Vaccines lead to the production of specialized cells and antibodies, which can efficiently and effectively target pathogens. There are several types of vaccines. Some consists of dead or inactivated pathogens or parts of them. The Pfizer-BioNTech and Moderna COVID-19 vaccines, approved by the FDA in December of 2020 consist of a piece of the the virus's genetic material (called mRNA). When injected, the body uses this genetic information to synthesize the spike protein of the virus, which stimulates an immune responses without causing infection.

Development and approval of a vaccine does not automatically eradicate a disease (figure 5.4.3.*g*), for the vaccine must be distributed to those who need it. Access is particularly challenging in developing countries, especially for low-income individuals there. In the case of tuberculosis, which remains widespread, a vaccine has been available for 100 years, but it is more effective in protecting severe forms of the disease in children than lung infections in adults. A new, more effective vaccine is likely needed to eradicate tuberculosis.

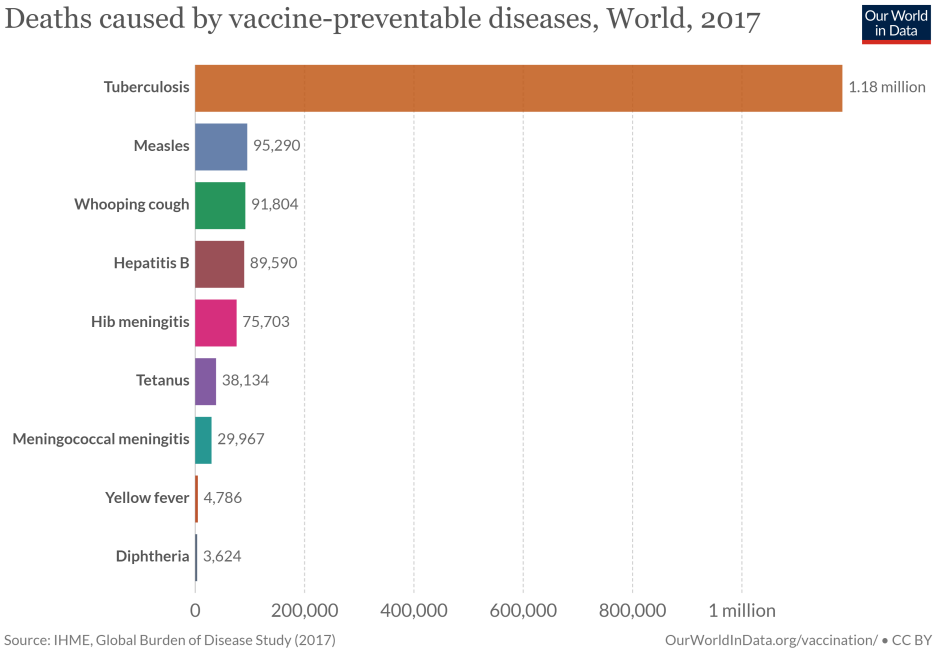


Figure 5.4.3.*g*: Global deaths in 2017 that were caused by diseases with existing vaccines. Tuberculosis was responsible for 1.18 million deaths followed by measles (95,290), whooping cough (91,804), hepatitis B (85,590), Hib meningitis (75,703), tetanus (38,134), meningococcal meningitis (29,967), yellow fever (4,786), and diphtheria (3,624). Image by Samantha Vanderslott and Bernadeta Dadonaite (2013) "Vaccination". Published online at OurWorldInData.org. (CC-BY).

New vaccines must be rigorously tested before they are approved. Not only are vaccines safe, but we would face spikes in deadly infectious diseases without them. The antivax movement has unfortunately spread false information questioning the safety of vaccines, but you can educate yourself with reliable information about the development, testing, and safety of vaccines here.

Antibiotic Resistance

Antibiotics and similar drugs, together called antimicrobial agents, have been used for the last 70 years to treat patients who have infectious diseases. Technically, **antibiotics** are substances that are produced by certain microorganisms (such as the fungus, *Penicillium*, or the bacterium *Streptomyces*) that inhibit the growth of other microorganisms. However, any drugs used to treat bacterial infections are commonly referred to as antibiotics. Neither of these definitions includes substances used to treat viral infections, which are called antivirals. Since the 1940s, the use of antiobitics has greatly reduced illness and death from infectious

diseases. However, these drugs have been used so widely and for so long that the infectious organisms the antibiotics are designed to kill have adapted to them, making the drugs less effective.

Overuse of antibiotics have allowed pathogens to evolve **antibiotic resistance**, which occurs when bacteria change in a way that reduces the effectiveness of drugs, chemicals, or other agents designed to cure or prevent infections (figure 5.4.3.*h*). Overuse takes many forms, such as the inclusion of antibiotics in household products or prescribing antibiotics to treat infections without first determining whether the infection is bacterial or viral (in which case, the antibiotic would not be effective). Similar to the evolution of pesticide resistance, widespread use of antibiotics favors individuals that happen to have gene versions conferring antibiotic resistance. These individuals have the best chance of surviving and reproducing, so much so, that an antibiotic-resistant strain evolves. The initial antibiotic is not longer effective against this strain, and it must be treated with alternative medications. For example, tuberculosis (TB) was nearly eliminated in most parts of the world, but drug-resistant strains have now reversed that trend.

A similar example of drug resistance has evolved in *Plasmodium*, the protozoa that cause malaria. The insecticide dichlorodiphenyltrichloroethane (DDT) was widely used to control malaria-carrying mosquito populations in tropical regions. However, after many years the mosquitoes developed a natural resistance to DDT and again spread the disease widely. Anti-malarial medicines were also over-prescribed, which allowed *Plasmodium* to become drug-resistant.

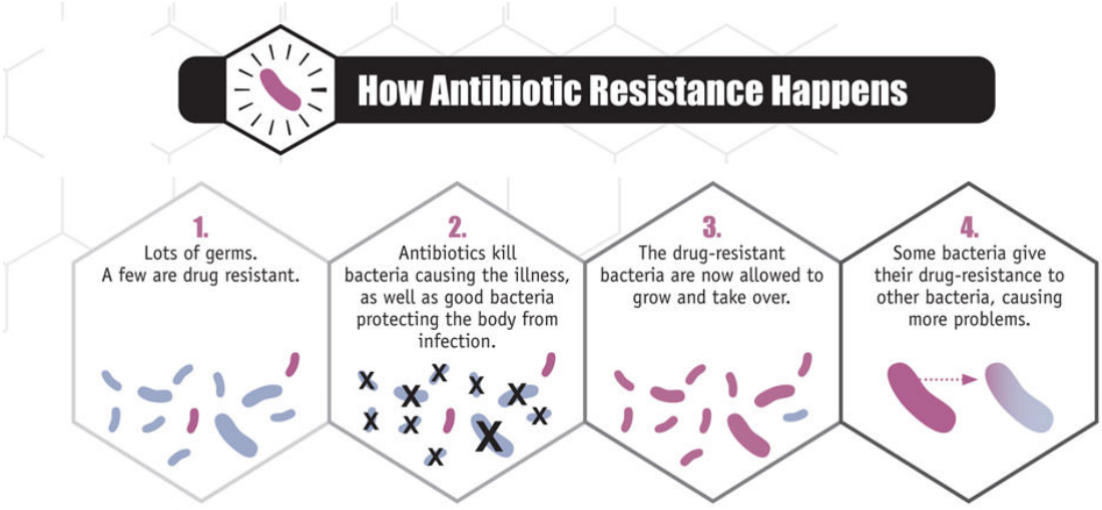


Figure 5.4.3.*h*: Antibiotic resistance is of increasing concern to medical professionals. (1) Initially, a variety of microbes (germs) are present, and a few are drug resistant. (2) When antibiotics are used, they kill both disease-causing microbes and beneficial ones. However, drug-resistant microbes remain. (3) The drug-resistant microbes multiple and can eventually produce a drug-resistant strain. (4) Drug-resistant bacteria can even pass their resistance genes to other bacteria, including those of different species.

New forms of antibiotic resistance can cross international boundaries and spread between continents with ease. Many forms of resistance spread with remarkable speed. Each year in the United States, at least 2 million people acquire serious infections with bacteria that are resistant to one or more of the antibiotics designed to treat those infections. At least 23,000 people die each year in the US as a direct result of these antibiotic-resistant infections. Many more die from other conditions that were complicated by an antibiotic-resistant infection. The use of antibiotics is the single most important factor leading to antibiotic resistance around the world. Antibiotics are among the most commonly prescribed drugs used in human medicine, but up to 50% of all the antibiotics prescribed for people are not needed or are not optimally effective as prescribed.

The video below shows an experiment done at Harvard Medical school, where they show bacteria adapting very quickly to apparently deadly conditions.



During recent years, there has been growing concern over methicillin-resistant *Staphylococcus aureus* (**MRSA),** a bacterium that is resistant to many antibiotics. In the community, most MRSA infections are skin infections. In medical facilities, MRSA causes life-threatening bloodstream infections, pneumonia and surgical site infections.

Attribution

Modified by Melissa Ha from the following sources:

- Environmental Health and Sustainable Agriculture from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Environmental Health and Child Survival: Epidemiology, Economics, Experiences. Environment and Development. 2008. World Bank. Washington, DC. CC-BY.
- Global Public Health from *Microbiology* by OpenStax (CC-BY). Access for free at openstax.org.
- Patel, V., D. Chisholm., T. Dua, R. Laxminarayan, and M. E. Medina-Mora, editors. 2015. *Mental, Neurological, and Substance Use Disorders*. Disease Control Priorities, third edition, volume 4. Washington, DC: World Bank. doi:10.1596/978-1-4648-0426-7. (CC-BY).
- Balkhair A. A. 2020. COVID-19 Pandemic: A New Chapter in the History of Infectious Diseases. *Oman medical journal*, *35*(2), e123. https://doi.org/10.5001/omj.2020.41 (CC-BY-NC).
- Tizaoui, K., Zidi, I., Lee, K. H., Ghayda, R. A., Hong, S. H., Li, H., Smith, L., Koyanagi, A., Jacob, L., Kronbichler, A., & Shin, J. I. (2020). Update of the current knowledge on genetics, evolution, immunopathogenesis, and transmission for coronavirus disease 19 (COVID-19). *International journal of biological sciences*, *16*(15), 2906–2923. https://doi.org/10.7150/ijbs.48812 (CC-BY).
- Ebola (Ebola Virus Disease). 2020. Centers for Disease Control and Prevention. Accessed 10 January 2021. (public domain)
- 2014-2016 Ebola Outbreak in West Africa. 2019. Centers for Disease Control and Prevention. Accessed 10 January 2021. (public domain)
- About HIV. 2020. Centers for Disease Control and Prevention. Accessed 10 January 2021. (public domain)
- HIV/AIDS at a glance. 2003. World Bank. Washington, DC. CC-BY.

This page titled 5.4.3: Infectious Diseases is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **15.3: Infectious Diseases** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

5.4.4: Environmental Toxicology

Environmental toxicology is the scientific study of the properties of **toxins**, chemicals that may cause damage to living organisms, and the health effects associated exposure to them (table 5.4.4.a). The field also involves managing these toxins and the protecting humans and ecosystems from them. **Toxicologists** are scientists who study the properties of toxins, and these damage-causing properties, are called **toxicity**. In other words, toxicologists evaluate chemical hazards.

Table 5.4.4.a**:** The 15 Highest Priority Toxins According to the Agency for Toxic Substances and Disease Registry (ATSDR) Substance Priority List in 2019. These substances are ranked based on their toxicity and chance of exposure at sites on the National Priorities List (NPL), where toxins have been spilled or otherwise released.

2019 Rank	Substance Name
1	Arsenic
2	Lead
3	Mercury
4	Vinyl chloride
5	Polychlorinated biphenyls (PCBs)
6	Benzene
7	Cadmium
8	Benzo[a]pyrene
9	Polycyclic aromatic hydrocarbons
10	Benzo[b]fluoranthene
11	Chloroform
12	Aroclor 1260
13	p,p'-DDT
14	Aroclor 1254
15	Dibenz[a,h]anthracene

Table modified from ATSDR/CDC (public domain).

What Forms do Chemicals Take?

Chemical substances can take a variety of forms. They can be in the form of solids, liquids, dusts, vapors, gases, fibers, mists and fumes (figure 5.4.4.a). The form a substance is in has a lot to do with how it gets into your body and what harm it can cause. A chemical can also change forms. For example, liquid solvents can evaporate and give off vapors that you can inhale. Sometimes chemicals are in a form that can't be seen or smelled, so they can't be easily detected.



Figure 5.4.4.*a*: Crystals of benzene under a layer of liquid benzene. They were formed inside a container exposed to temperatures below 5.5°C. Their smoothness is the result of taking the photograph some time after bringing the container inside an atmosphere at room temperature, where they began to melt. Image and caption by Endimion17 (CC-BY-SA).

Routes of Exposure to Chemicals

In order to cause health problems, chemicals must enter your body. Once chemicals have entered your body, some can move into your bloodstream and reach internal "target" organs, such as the lungs, liver, kidneys, or nervous system. There are three main **routes of exposure**, or ways a chemical can get into your body (figure 5.4.4.*b*).

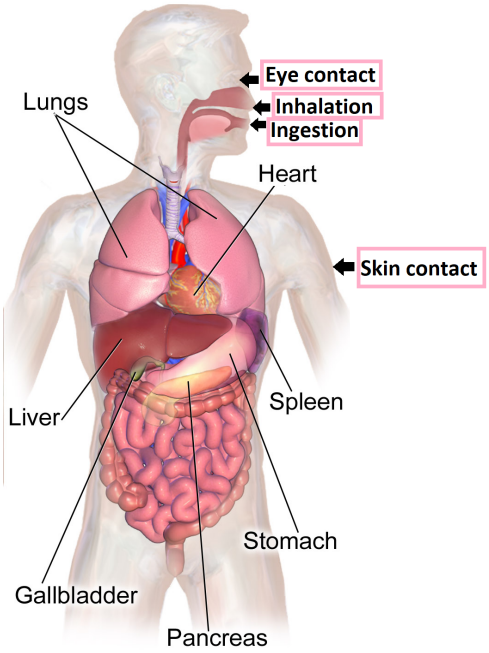


Figure 5.4.4.*b*: The routes of exposure through which toxins can enter the body (inhalation, ingestion, and skin or eye contact) are boxed in pink. Image by BruceBlaus (CC-BY-SA).

The first route is **inhalation**, which results from breathing in chemical gases, mists, or dusts that are in the air. Because the air sacs in the lungs are structure for gas exchange, consisting of only a single layer of thin cells, inhaled substances can rapidly pass from the air sacs into capillaries and enter the bloodstream. Inhaled toxins can also cause local damage to the mouth, airways, and lungs.

The second route is **ingestion**, swallowing that occurs when chemicals have spilled or settled onto food, beverages, cigarettes, beards, or hands. Toxins can then cause local damage to the digestive tract and may be absorbed, most often through the intestines, into the blood stream.

The third route is **skin or eye contact**. When chemicals directly touch the skin or get in the eyes, they can cause localized damage or be absorbed through the skin into the bloodstream. Because the skin consists of many cell layers that are reinforced with protective compounds, it is more difficult for toxins to enter the body through the skin compared to inhalation and ingestion. However, toxins can easily enter through damaged skin, and some substances can be absorbed through intact skin. Because the eyes have a rich blood supply, toxins can ready enter the bloodstream through eye contact.

What Factors Impact a Chemical's Safety?

Why are some chemicals more harmful than other? Many factors are considered when evaluating a chemical's safety including potency, persistence, solubility, bioaccumulation and biomagnification. **Potency** refers to the amount of a chemical needed to cause harm. The more potent a toxin, the lower the concentration needed to cause harm. **Persistence** refers to how long a substance takes to break down. Persistent chemicals are of greater concern because they remain in the environment (or even in organisms) for long time periods. **Solubility** refers to whether the chemical dissolves in certain solvents, such as water or fat. Generally, fat-soluble (lipid-soluble) toxins are more dangerous because they can accumulate in fatty tissues (see below) whereas water-soluble toxins could be more easily flushed out of the body. Furthermore, fat-soluble toxins are more easily absorbed by the body.

Bioaccumulation

Bioaccumulation is the buildup of chemicals in an organism's tissues over its lifetime. While bioaccumulated toxins are commonly fat soluble, such as DDT and PCBs, water-soluble toxins, such as inorganic forms of heavy metals can also bioaccumulate. For example, lead accumulates in the teeth and bone, and mercury can accumulate in the kidneys and brain.

Biomagnification

Biomagnification is the increasing concentration of toxins in organisms at each successive trophic level. When organisms with bioaccumulated toxins are consumed, the toxins are transferred to their predators (figure 5.4.4.*d*). Biomagnification explains why some fish species high on the food chain contain high concentrations of mercury and cadmium, another heavy metal.

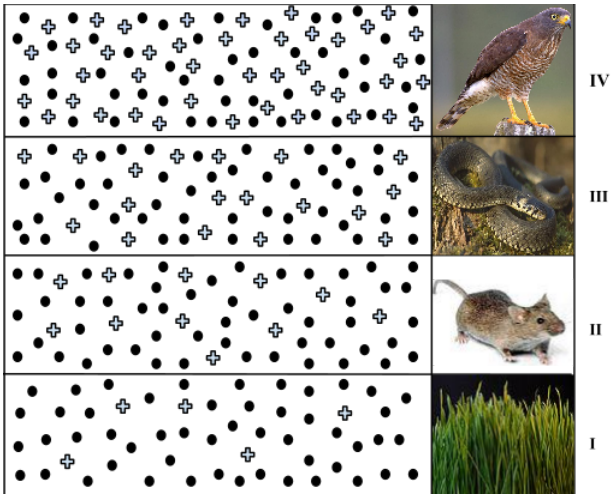


Figure 5.4.4.*c*: Biomagnification. The biomagnified toxin becomes concentrated in the tissues of organisms representing four successive trophic levels in a food chain. Most of the toxin absorbed by primary producers remains in their bodies at that first trophic level. Through biomagnification, the concentration of the toxin (crosses) increases higher up the food chain. (Black dots represent other molecules.) Organisms at the top thus have a higher tissue concentration of toxins than lower levels. Trophic level I represents the primary producers; trophic level II represents the primary consumers; trophic level III represents the secondary consumers; and trophic level IV represents the tertiary consumers. Image by Sballesteros15 (CC-BY-SA).

In addition to high persistence, fat solubility, and bioaccumulation, biomagnification explains why the now-banned insecticide dichlorodiphenyltrichloroethane (DDT) caused so much damage. Producers absorbed DDT and passed it on to successive levels of consumers at increasingly higher rates. For example, spraying a marsh to control mosquitoes will cause trace amounts of DDT to accumulate in the cells of microscopic aquatic organisms, the plankton, in the marsh. In feeding on the plankton, filter-feeders, like clams and some fish, harvest DDT as well as food. (Concentrations of DDT 10 times greater than those in the plankton have been measured in clams.) The process of concentration goes right on up the food chain from one trophic level to the next. Gulls, which feed on clams, may accumulate DDT to 40 or more times the concentration in their prey. This represents a

400-fold increase in concentration along the length of this short food chain. Ultimately, apex predators at the top of the food chain such as Bald Eagles, pelicans, falcons, and ospreys fed on contaminated fish, reaching dangerous DDT levels.

Another substances that biomagnifies is polychlorinated biphenyl (PCB). The National Atmospheric and Oceanic Administration (NOAA) studied biomagnification of PCB in the Saginaw Bay of Lake Huron of the North American Great Lakes (figure 5.4.4.*d*). Concentrations of PCB increased from the producers of the ecosystem (phytoplankton) through the different trophic levels of fish species. The apex predator, the walleye, had more than four times the amount of PCBs compared to phytoplankton. Also, research found that birds that eat these fish may have PCB levels that are at least ten times higher than those found in the lake fish. This aquatic ecosystem offered an ideal opportunity to study biomagnification because PCB generally exists at low concentrations in this environment, but apex predators accumulated very high concentrations of the toxin.

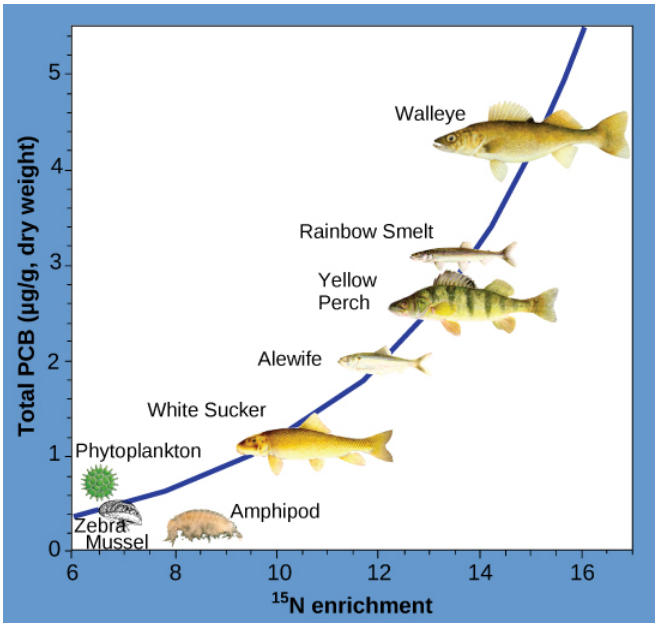


Figure 5.4.4.*d*: Polychlorinated biphenyl (PCB) concentrations found at the various trophic levels in the Saginaw Bay ecosystem of Lake Huron. Numbers on the x-axis reflect enrichment with heavy isotopes of nitrogen (^{15}N), which is a marker for increasing trophic level. Notice that the fish in the higher trophic levels accumulate more PCBs than those in lower trophic levels. The organisms in order of ^{15}nitrogen enrichment are phytoplankton, zebra mussel, amphipod, white sucker, alewife, yellow perch, rainbow smelt, and walleye. (credit: Patricia Van Hoof, NOAA, GLERL)

Persistent Organic Pollutants (POPs)

Persistent organic pollutants (POPs) are a group of organic chemicals that pose risks to human health and ecosystems. Examples include the pesticide dichlorodiphenyltrichloroethane (DDT) and the industrial chemicals polychlorinated biphenyls (PCBs) and per- and polyfluoroalkyl substances (PFAS). The contaminant in Agent Orange (2,3,7,8-tetrachlorodibenzo-p-dioxin; TCDD) is another POP (see The Scientific Method). Persistent organic pollutants have the following three characteristics:

- **Persistent**: POPs are chemicals that last a long time in the environment. Some may resist breakdown for years and even decades while others could potentially break down into other toxic substances.
- **Bioaccumulative**: POPs can accumulate in animals and humans, usually in fatty tissues and largely from the food they consume. As these compounds move up the food chain, they concentrate to levels that could be thousands of times higher than acceptable limits.
- **Toxic**: POPs can cause a wide range of health effects in humans, wildlife and fish. They have been linked to effects on the nervous system, reproductive and developmental problems, suppression of the immune system, cancer, and endocrine disruption. The deliberate production and use of most POPs has been banned around the world, with some exemptions made for human health considerations (for example, DDT for malaria control) and/or in very specific cases where alternative chemicals have not been identified. However, the unintended production and/or the current use of some POPs continue to be an issue of global concern. Even though most POPs have not been manufactured or used for decades, they

continue to be present in the environment and thus potentially harmful. The same properties that originally made them so effective, particularly their stability, make them difficult to eradicate from the environment.

The relationship between exposure to environmental contaminants such as POPs and human health is complex. There is mounting evidence that these persistent, bioaccumulative and toxic chemicals (PBTs) cause long-term harm to human health and the environment. Drawing a direct link, however, between exposure to these chemicals and health effects is complicated, particularly since humans are exposed on a daily basis to many different environmental contaminants through the air they breathe, the water they drink, and the food they eat. Numerous studies link POPs with a number of adverse effects in humans. These include effects on the nervous system, problems related to reproduction and development, cancer, and genetic impacts. Moreover, there is mounting public concern over the environmental contaminants that mimic hormones in the human body (**endocrine disruptors**).

Through atmospheric processes, they are deposited onto land or into water ecosystems where they accumulate and potentially cause damage. From these ecosystems, they evaporate, again entering the atmosphere, typically traveling from warmer temperatures toward cooler regions. They condense out of the atmosphere whenever the temperature drops, eventually reaching highest concentrations in circumpolar countries. Through these processes, POPs can move thousands of kilometers from their original source of release in a cycle that may last decades.

As with humans, animals are exposed to POPs in the environment through air, water and food. POPs can remain in sediments for years, where bottom-dwelling creatures consume them and who are then eaten by larger fish. Because tissue concentrations can biomagnify at each level of the food chain, top predators, including whales, seals, polar bears, birds of prey, tuna, swordfish and bass may have a million times greater concentrations of POPs than the water itself. Once POPs are released into the environment, they may be transported within a specific region and across international boundaries transferring among air, water, and land.

While generally banned or restricted (figure $5.4.4.e$), POPs make their way into and throughout the environment on a daily basis through a cycle of long-range air transport and deposition called the "**grasshopper effect.**"The "grasshopper" processes begin with the release of POPs into the environment. When POPs enter the atmosphere, they can be carried with wind currents, sometimes for long distances.

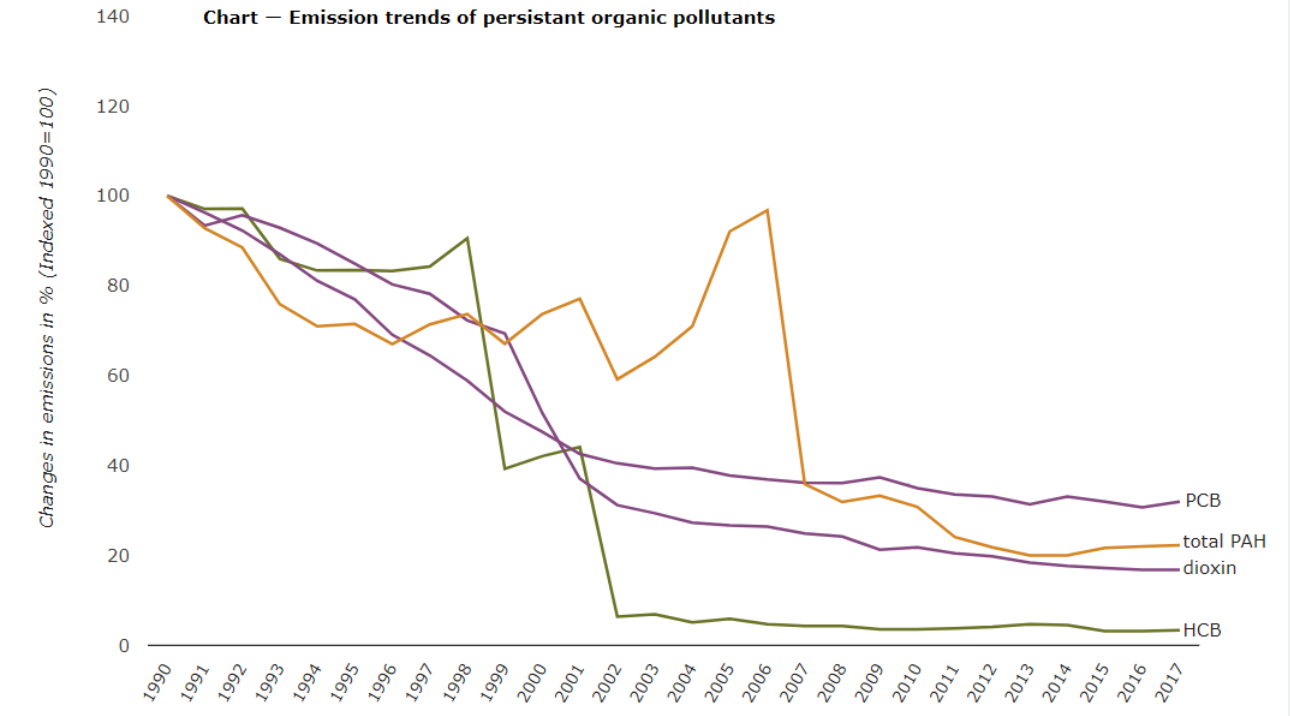


Figure 5.4.4.*e*: Emissions of several persistent organic pollutants (POPs) by European Union countries has decreased over the years. These include hexchlorobenze (HCB), polychlorinated biphenyls (PCBs), dioxins (such TCDD), and total polycyclic aromatic hydrocarbons (PAHs), However, these compounds remain in the environment for long periods of time. Image modified from European Environment Agency of the European Union (CC-BY)

What Are the Health Effects of Toxins?

A toxin's effect is determined by many factors. Firstly, the same toxin and the same concentration can affect individuals differently depending on age, overall health, genetics, gender, and other factors. For example, young children are especially susceptible to heavy metals and bisphenol A (BPA). Additionally, as many toxins are processed by the liver, and liver function decreases with age, elderly individuals are more susceptible to certain toxins. Similarly, an generally health individual is likely to withstand exposure toxins better than someone facing other health concerns. Regarding genetics, some individuals may have versions of genes that make them more susceptible or resistant to certain toxins. In addition to individual factors, the length of exposure, presence of other toxins, and concentration (dosage) all impact toxicity

Acute vs. Chronic Effects

An **acute effect** of a toxin is one that occurs rapidly after exposure to a large amount of that substance. A **chronic effect** of a contaminant results from exposure to small amounts of a substance over a long period of time. In such a case, the effect may not be immediately obvious. Chronic effects are difficult to measure, as the effects may not be seen for years. Long-term exposure to cigarette smoking, low-level radiation exposure, and moderate alcohol use are all thought to produce chronic effects.

Toxicological Interactions

Exposure to multiple chemicals simultaneously can result in a variety of effects, or **chemical interactions** (figure 5.4.4.*f*). These can apply to any chemicals that affect the body, including drugs and toxins (in the latter case, they are called **toxicological interactions**.) If the effect of the two chemicals combined is the sum of their individual effects, the chemical interaction is called an **additive effect**. Suppose that drug A and drug B have the same effect on the body (for example, each increasing heart rate by five beats per minute). An additive effect would occur if taking both drugs together increased the heart rate by 10 beats per minute. For example, toluene and xylene, which are both found in solvents and paints have an additive effect. Each causes eye, nose, and throat irritation, dizziness, headaches, and confusion. A study found that their combined effects on memory, cognitive function, and coordination was roughly equal to the sum of their individual effects.

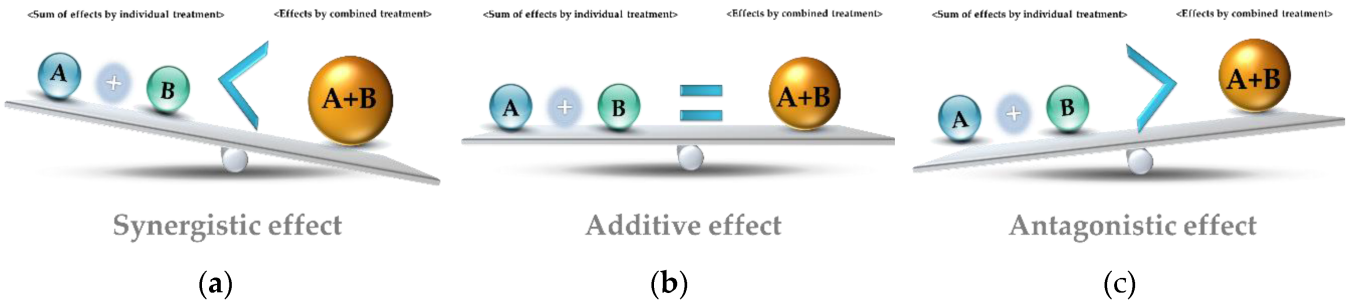


Figure 5.4.4.*f*: Chemicals interactions in the body may have a synergistic (a), additive (b), and antagonistic (c) effect. For each, chemicals A and B are on one side of a "seesaw", and their combined effect (A+B) is on the other side. In the synergistic effect, the combined effect exceeds the sum of the effects of individual treatments. The opposite is true for the antagonistic effect. In the additive effect, the combined effect equals the sum of effects by individual treatments. Image by Tae Jin Cho et al., MDPI, Basel, Switzerland (CC-BY).

In contrast, if taking both drugs A and B simultaneously increased heart rate by less than 10 beats per minute (less than their sum), their interaction would be considered an **antagonistic effect**, indicating that they interfere with each other. For example, ethanol (found in alcoholic beverages) and methanol (wood alcohol) both have toxic effects, but methanol is more immediately damaging. When methanol is ingested, they body converts it to dangerous compounds that cannot be easily removed from the body. Ethanol can used to treat methanol poisoning because it blocks the enzymes that facilitate these reactions, providing the body an opportunity to eliminate the methanol.

If combining drugs A and B increased heart rate more than 10 beats per minute, their interaction would be a **synergistic effect**; that is, their combined effect was greater than the sum of their individual effects. For example, smoking combined with asbestos exposure have a synergistic effect in causing lung cancer.

The Dose-Response Curve

For centuries, scientists have known that just about any substance is toxic in sufficient quantities. It is a thus common saying among toxicologists that "the dose makes the poison". In fact, too high a dosage (amount) of an otherwise helpful drug can cause negative health effects or even death. Similarly, small amounts of selenium are required by living organisms for proper functioning, but large amounts may cause cancer.

The effect of a certain chemical on an individual depends on the dose of the chemical. This relationship is often illustrated by a **dose-response curve**, which shows the relationship between dose and the response of the individual. **Lethal doses** in humans have been determined for many substances from information gathered from records of homicides, accidental poisonings, testing on animals, and experiments on cell cultures. A dose that is lethal to 50% of a population of test animals is called the **lethal dose-50%**, or **LD_{50}** (Table 5.4.4.*b*). Determination of the LD_{50} is required for new synthetic chemicals in order to give a measure of their toxicity. Because a single conventional LD_{50} test may kill as many as 100 animals, the United States and other members of the Organization for Economic Cooperation and Development agreed in December 2000 to phase out the LD_{50} test in favor of alternatives that greatly reduce (or even eliminate) deaths of the test animals.

Table 5.4.4.*b*: The LD_{50} values for some insecticides. In each case, the chemical was fed to laboratory rats. Note that the lower the LD_{50}, the more toxic the chemical.

Chemical	Category	Oral LD_{50} in Rats (mg/kg)
Aldicarb ("Temik")	Carbamate	1
Carbaryl ("Sevin")	Carbamate	307
DDT	Chlorinated hydrocarbon	87
Dieldrin	Chlorinated hydrocarbon	40
Diflubenzuron ("Dimilin")	Chitin inhibitor	10,000
Malathion	Organophosphate	885
Methoprene	JH mimic	34,600
Methoxychlor	Chlorinated hydrocarbon	5,000
Parathion	Organophosphate	3
Piperonyl butoxide	Synergist	7,500
Pyrethrins	Plant extract	200
Rotenone	Plant extract	60

A dose that causes 50% of a population to exhibit any significant response, whether therapeutic or harmful (hair loss, stunted development, etc.), is referred to as the **effective dose-50%**, or **ED_{50}** (figure 5.4.4.*h*). Some toxins have a threshold dose below which there is no apparent effect on the exposed population, called the **no-observed-adverse-effect level** (**NOAEL;** figure 5.4.4.*i*). The lowest dose at which any negative effect is apparent is called the **lowest-observed-adverse-effect level** (**LOAEL**). Between NOAEL and LOAEL, there may be a noticeable but harmless effect.

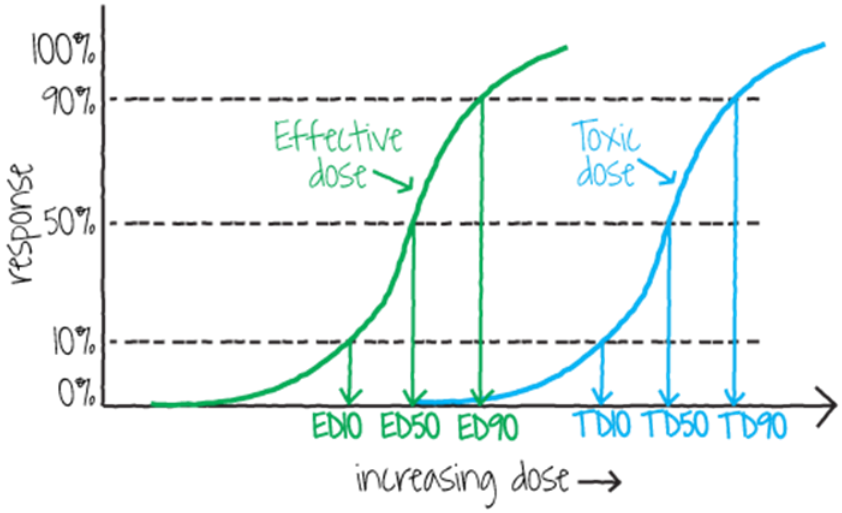


Figure 5.4.4.*h*: A dose-response curve shows the percentage of individuals demonstrating a response to a drug based on dosage. The **effective dose-50%** (ED_{50}) is the dosage at which half of the population demonstrates an effect. The toxic dose-50% (TD_{50}) is the dosage at which half of the population is poisoned by the drug. The **lethal dose-50%** (LD_{50}), the dosage at which half of the population dies, could be illustrated by a third curve to the right. Image by U.S. National Library of Medicine (public domain).

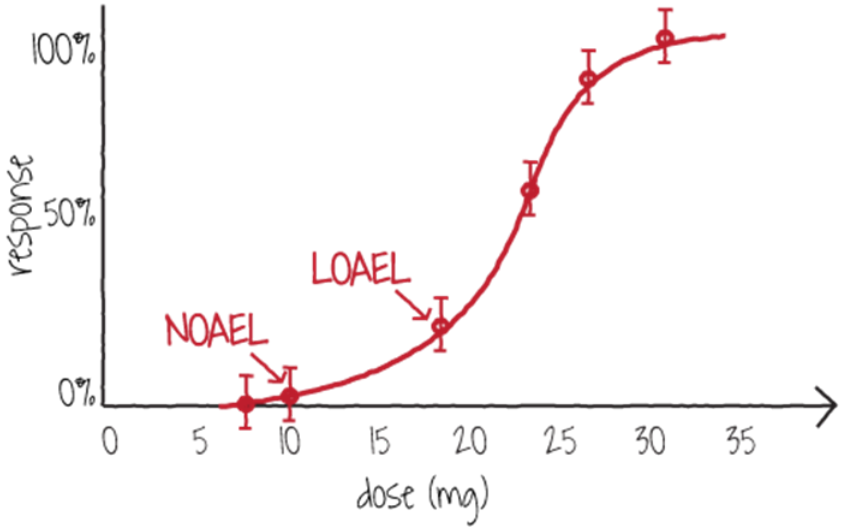


Figure 5.4.4.*i*: A dose-response curve illustrating the **no-observed-adverse-effect level** (**NOAEL**) and **lowest-observed-adverse-effect level** (**LOAEL**). Image by U.S. National Library of Medicine (public domain).

Human Health, Environmental, and Economic Effects of Pesticide Use in Potato Production in Ecuador

The International Potato Center (CIP) conducted an interdisciplinary and inter-institutional research intervention project dealing with pesticide impacts on agricultural production, human health, and the environment in Carchi, Ecuador. Carchi is the most important potato-growing area in Ecuador, where smallholder farmers dominate production. They use tremendous amounts of pesticides for the control of the Andean potato weevil and the late blight fungus. Virtually all farmers apply highly hazardous pesticides (classified by the World Health Organization as class Ib) using hand pump backpack sprayers (figure 5.4.4.*g*). The LD_{50} for class Ib pesticides is 5-50 or 20-200 mg/kg for oral ingestion in solid or liquid form, respectively. For dermal (skin) exposure, the LD_{50} is 10-100 or 40-400 mg/kg for solids and liquids, respectively. The LD_{50} class III pesticides, which are considered only slightly hazardous, is at least 10 times greater that that for class Ib pesticides.



Figure 5.4.4.*g*: This Vietnam farmer sprays pesticides using a backpack sprayer with little personal protection, similar to what is used by Ecuador potato farmers. Image by Roy Bateman at English Wikipedia (CC-BY-SA).

The study found that the health problems caused by pesticides are severe and are affecting a high percentage of the rural population. Despite the existence of technology and policy solutions, government policies continue to promote the use of pesticides. The study conclusions concurred with those by the pesticide industry, “that any company that could not ensure the safe use of highly toxic pesticides should remove them from the market and that it is almost impossible to achieve safe use of highly toxic pesticides among small farmers in developing countries.”

Source: Yanggen et al. 2003.

Precautionary Principle

Determining a safe dosage from a dose-response curve employs the **precautionary principle**, which put simply, embodies the phrase “It’s better safe than sorry.” Actions that follow the precautionary principle allow margin to ensure safety in the case that a toxin or drug is later found to have negative effect at a lower dosage than first detected. The safe dosage is often set to 1% or even 0.1% of the NOAEL.

The precautionary principle is sometimes applied to other components of environmental toxicology. For example in the European Union (EU), manufacturers must demonstrate the safety of their product before it is sold. While this is also required in the U.S. for chemicals that have not been used before, there is no such rule for existing products. It is instead the responsibility of Environmental Protection Agency to demonstrate that products already on the market are unsafe before banning them. In summary, the precautionary principle is applied to regulating potential toxins in products in both the EU and U.S.; however, the EU's applies the precautionary principle more broadly. The EU errors on the side of caution, potentially banning chemicals that are harmless until they are proven safe, but the U.S. risks exposure to potential toxins before their safety is determined.

Attribution

Modified by Melissa Ha from the following sources:

- Persistent Organic Pollutants: Backyards to Borders. 2004. World Bank. (licensed under CC-BY)
- *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
  - Pests and Pesticides
  - Environmental Toxicology
  - The Impacts of Environmental Conditions
  - Environmental Health
  - Energy Flow through Ecosystems
- Biomagnification, Insecticides, and the LD50 test from *Biology* by John W. Kimball (licensed under CC-BY)

This page titled 5.4.4: Environmental Toxicology is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **15.4: Environmental Toxicology** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

5.4.5: Environmental Hazard Reduction

Public Health Organizations

A large number of international programs and agencies are involved in efforts to promote global public health. Among their goals are developing infrastructure in health care, public sanitation, and public health capacity; monitoring infectious disease occurrences around the world; coordinating communications between national public health agencies in various countries; and coordinating international responses to major health crises. In large part, these international efforts are necessary because disease-causing microorganisms know no national boundaries.

International public health issues are coordinated by the **World Health Organization** (WHO), an agency of the United Nations. Of its roughly $4 billion budget for 2015–16, about $1 billion was funded by member states and the remaining $3 billion by voluntary contributions. In addition to monitoring and reporting on infectious disease, WHO also develops and implements strategies for their control and prevention. WHO has had a number of successful international public health campaigns. For example, its vaccination program against smallpox, begun in the mid-1960s, resulted in the global eradication of the disease by 1980. WHO continues to be involved in infectious disease control, primarily in the developing world, with programs targeting malaria, HIV/AIDS, and tuberculosis, among others. It also runs programs to reduce illness and mortality that occur as a result of violence, accidents, lifestyle-associated illnesses such as diabetes, and poor health-care infrastructure.

WHO maintains a global alert and response system that coordinates information from member nations. In the event of a public health emergency or epidemic, it provides logistical support and coordinates international response to the emergency. The United States contributes to this effort through the **Centers for Disease Control and Prevention** (CDC), an agency of the Department of Health and Human Services (figure 5.4.5.a). The CDC carries out international monitoring and public health efforts, mainly in the service of protecting US public health in an increasingly connected world. Similarly, the European Union maintains a Health Security Committee that monitors disease outbreaks within its member countries and internationally, coordinating with WHO.



Figure 5.4.5.a: The Centers for Disease Control and Prevention is charged with protecting the public from disease and injury. Image by CDC (public domain).

One way that the CDC carries out this mission is by overseeing the National Notifiable Disease Surveillance System (NNDSS) in cooperation with regional, state, and territorial public health departments. The NNDSS monitors diseases considered to be of public health importance on a national scale. Such diseases are called notifiable diseases or reportable diseases because all cases must be reported to the CDC. A physician treating a patient with a notifiable disease is legally required to submit a report on the case. Notifiable diseases include HIV infection, measles, West Nile virus infections, and many others. Some states have their own lists of notifiable diseases that include diseases beyond those on the CDC's list.

Notifiable diseases are tracked by epidemiological studies and the data is used to inform health-care providers and the public about possible risks. The CDC publishes the *Morbidity and Mortality Weekly Report* (*MMWR*), which provides physicians and health-care workers with updates on public health issues and the latest data pertaining to notifiable diseases. Table 5.4.5.a is an example of the kind of data contained in the *MMWR*.

Table 5.4.5.a**:** Incidence of Four Notifiable Diseases in the United States, Week Ending January 2, 2016

Disease	Current Week (Jan 2, 2016)	Median of Previous 52 Weeks	Maximum of Previous 52 Weeks	Cumulative Cases 2015
Campylobacteriosis	406	869	1,385	46,618
Chlamydia trachomatis infection	11,024	28,562	31,089	1,425,303
Giardiasis	115	230	335	11,870
Gonorrhea	3,207	7,155	8,283	369,926

The current *Morbidity and Mortality Weekly Report* is available online.

Strategies for Reducing Environmental Hazards

Providing access to clean water is an important strategy for reducing environmental hazards. Because lack of clean water is responsible for diarrheal and other infectious diseases, particular in developing countries, access to clean water limits disease spread (a biological hazard). Furthermore, clean water reduces exposure to toxins (chemical hazards). This can be achieved through digging wells and establishing water treatment facilities. Improving sanitation and hygiene (for example, by supplying latrines, [figure 5.4.5.*b*] and handwashing stations) further reduces the spread of infectious disease.



Figure 5.4.5.*b*: When adequate plumbing is not available, latrines, such as this pit latrine, prevent sewage from polluting water supplies. Image by SuSanA Secretariat (CC-BY).

In Sub-Saharan Africa and South Asia, the use of pipe filters helped eradicate Guinea worm disease, which is caused by a parasitic worm. People are infected by this disease when they drink water contaminated by small animals called copepods that house the larvae. Once in the body, the larvae mature in the abdomen, and adult worms eventually (and painfully) exit through the skin. If an infected person enters a body of water at this stage, adult worms release more larvae into the water, continuing their life cycle. Pipe filters are straw-like structures that contain openings small enough to allow the passage of water but not copepods, preventing infection.

Another strategy to limit exposure to biological hazards is to reduce exposure to disease vectors. Depending on the specific vector and disease this could involve removing standing water (which facilitates mosquito reproduction), application of pesticides, or use of netting (figure 5.4.5.*c*). Biological control is also slowly emerging in vector control in public health and in areas that for a long time mainly focused on chemical vector control of the *Anopheles* mosquito (the vector of malaria) and the black fly (the vector of river blindness, caused by a parasitic worm). The release of sterile males has been used to control the tsetse fly, the vector of African sleeping sickness. Vaccinating animals that harbor diseases (reservoirs) can also limit the chance of infection.



Figure 5.4.5.*c*: A net excludes mosquitos, which are disease vectors. The *Anopheles* mosquito is only active at night, and thus sleeping inside of a net can limit exposure to malaria. Image by Presidents Mosquito Initiative a component of the U.S. Government's Global Health Initiative (public domain).

Because air pollution exposes individuals to toxins, limiting air pollution is another means of reducing environmental hazards. This can involve energy conservation, using clean energy, such as solar or wind (rather than burning fossil fuels), enforcing air pollution standards on industry, and implementing pollution-reducing technologies, at an industry or household scale. A main source of indoor air pollution is cooking using a fire indoors. Solar ovens provide a pollution-free alternative for people who do not have access to electricity or gas for cooking (figure 5.4.5.*d*).



Figure 5.4.5.*d*: Solar ovens are a pollution-free option for cooking food if electricity or gas is not available. Image by SUN OVEN (public domian).

The public have an important role in environmental hazard reduction. Firstly, the public must engage in behaviors to prevent disease spread and minimize contact with toxins, and public health education is critical to this. In the case of Guinea worm disease, individuals learned how to use pipe filters and shy they must avoid submerging wounds from the worms in bodies of water to limit disease spread. In the case of COVID-19, individuals are learning to reduce disease transmission by employing frequent handwashing, mask wearing, and social distancing (figure 5.4.5.*e*).

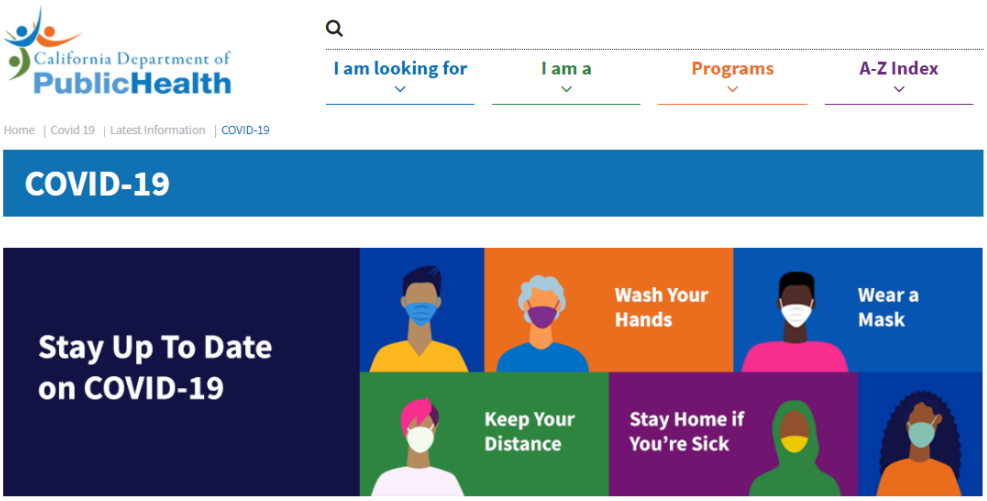


Figure 5.4.5.*e*: The California Department of Public Health website includes guidelines for limiting disease spread. The text reads "Stay Up To Date on COVID-19. Wash Your Hands. Wear a Mask. Keep Your Distance. Stay Home if You're Sick." Image from the CDPH website (public domain).

The public can also support policies that reduce exposure to environmental hazards. For example, many states have laws that reduce exposure to secondhand smoke. The Toxic Substances Control Act (TSCA) bans or regulates harmful chemicals, such as asbestos and heavy metals (figure 5.4.5.*f*)..



Figure 5.4.5.*f*: Asbestos removal from a building in the United Kingdom (UK). While asbestos is fully banned in the UK, only certain uses of asbestos are banned in the United States under the Toxic Substances Control Act . Image by Fevs101 (CC-BY-SA).

Many strategies for reducing environmental health hazards directly align with the United Nations Sustainable Development Goals, particularly goals 3 (Good Health and Well-Being) and 6 (Clean Water and Sanitation).

Attribution

Modified by Melissa Ha from the following sources:

- Disease and Epidemiology from *Microbiology* by OpenStax (licensed under CC-BY)
- Sustainable Agriculture from *Environmental Biology* by Mathew R. Fisher (licensed under CC-BY)

This page titled 5.4.5: Environmental Hazard Reduction is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **15.5: Environmental Hazard Reduction** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

https://bio.libretexts.org/@go/page/109193

5.4.6: Data Dive- Cholera Cases Worldwide

Overview

The World Health Organization (WHO) seeks to promote health and safety for humanity, especially the most vulnerable. What this means is they do everything from basic primary care, to research, to data inventory, to health education, and to many other imperative health initiatives worldwide. When it comes to data, they keep up to date records for chronic and acute diseases, especially those that can cause global/local health emergencies. Generally, the WHO tracks the number of cases (infection rates) and case fatalities for each disease. However, there are many other statistics they can analyze depending on the level of accuracy and detail of the data they receive. They can also break up or group their statistics by country or by continent (as you see below). In addition, the WHO can be a great database for researchers around the globe. Below is a graph created by WHO data on cholera from a 2019 publication about cholera epidemiology:

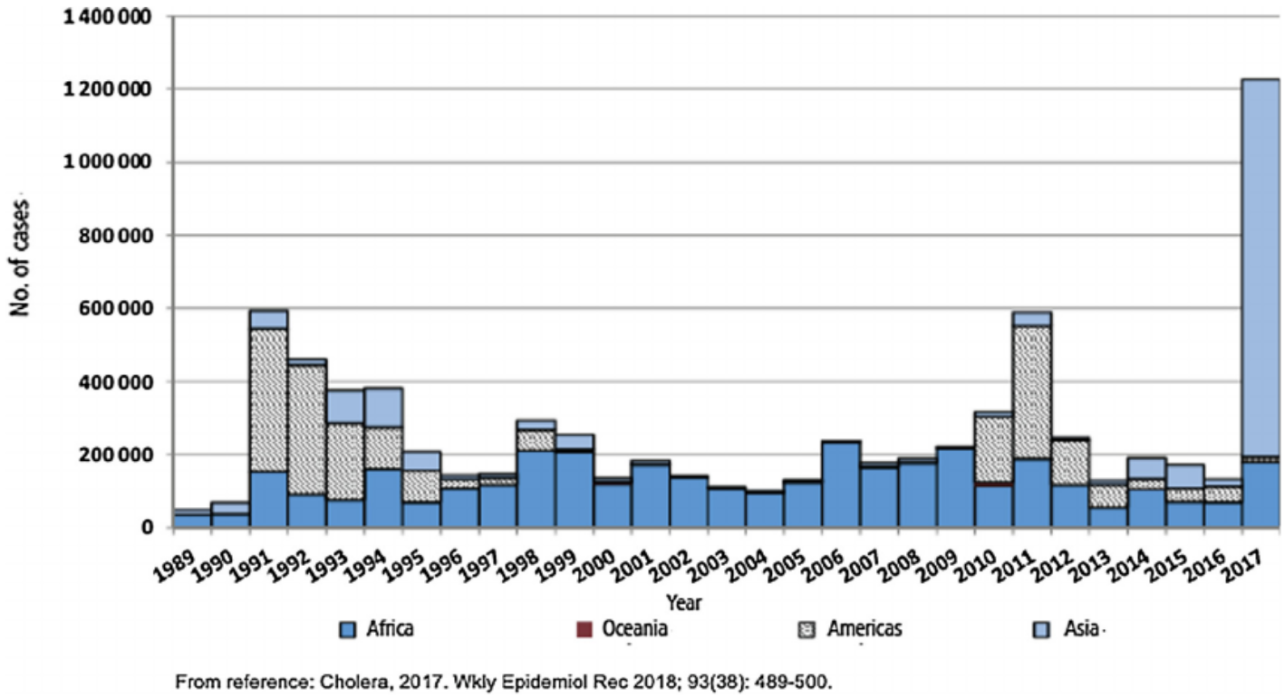


Figure 5.4.6.a: Number of cholera cases by world region for 20+ year period. Graph by Deen J, Mengel MA, Clemens JD 2019 (CC-BY4.0)

Questions

1. What kind of graph is this?
2. What is the independent (explanatory) variable and the dependent (response) variable?
3. Which continent has the most consistent cholera cases?
4. Describe the patterns you see in cholera cases for the Americas and Asia from 1989 to 2017.
5. In what ways do you think the WHO uses data observed in the graph above to make choices into the future? Provide at least two ideas/examples.

Raw Data For Above Graph(s)

Raw data table link for above graph. (CC-BY4.0)

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 5.4.6: Data Dive- Cholera Cases Worldwide is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **15.6: Data Dive- Cholera Cases Worldwide** by Melissa Ha and Rachel Schleiger has no license indicated.

5.4.7: Review

Summary

After completing this chapter you should be able to...

- Compare traditional and modern environmental hazards.
- Provide examples of biological, chemical, and physical hazards.
- Distinguish between morbidity and mortality.
- Distinguish between disease prevalence and incidence.
- Compare sporadic, endemic, epidemic, and pandemic patterns of incidence.
- Compare contact, vector, and vehicle modes of transmission and provide examples of how environmental degradation has increased transmission.
- Identify examples of emerging and reemerging infectious diseases.
- Briefly explain how vaccines work and their role in eradicating infectious diseases.
- Describe how antibiotic resistance evolves.
- Identify the three main routes of exposure to chemicals.
- Identify and define five factors that impact a chemical's safety.
- Compare additive, antagonistic, and synergistic toxicological interactions.
- Explain how the lethal dose-50% is determined from a dose-response curve.
- Provide examples of strategies for reducing environmental hazards.

Environmental health focuses on how the natural and human-built surroundings as well as behaviors affect human well-being. **Traditional hazards** are related to poverty and mostly affect low-income people and those in developing countries whereas **modern hazards** result from technological development and prevail in industrialized countries. Environmental health hazards are classified as biological (infectious diseases), chemical (exposure to toxins), or physical.

Epidemiology is the science underlying public health. **Morbidity** means being in a state of illness, whereas **mortality** refers to death; both **morbidity rates** and **mortality rates** are of interest to epidemiologists. **Incidence** is the number of new cases (morbidity or mortality), usually expressed as a proportion, during a specified time period; **prevalence** is the total number affected in the population, again usually expressed as a proportion. **Sporadic diseases** only occur rarely and largely without a geographic focus. **Endemic diseases** occur at a constant (and often low) level within a population. **Epidemic diseases** and **pandemic diseases** occur when an outbreak occurs on a significantly larger than expected level, either locally or globally, respectively. Diseases may be spread through **contact transmission** (direct physical interaction or indirect interaction through shared objects), **vehicle transmission** (through water, food, or air), or **vector transmission** (by an animal). Ecosystem degradation can exacerbate disease spread through creating conditions where vectors, pathogens, or reserviors thrive.

Emerging and reemerging diseases have been defined as infectious diseases of humans whose occurrence during the past two decades has substantially increased or threatens to increase in the near future relative to populations affected, geographic distribution, or magnitude of impacts. **Vaccination** has been used to successfully eradicate several infectious diseases by stimulating a specific immune response. For remaining diseases that are commonly treated with medications, **antibiotic resistance** continues to be a global problem. New forms of antibiotic resistance can cross international boundaries and spread between continents.

Environmental toxicology is the scientific study of the health effects associated with exposure to toxins. The three mains routes of exposure to toxins are inhalation, ingestion, and skin or eye contact. Potency, persistence, solubility, bioaccumulation, and biomagnification all affect a chemical's safety. **Bioaccumulation** refers to the buildup of a toxin over an individual's lifetime whereas **biomagnification** occurs to increasing concentrations of a toxin as it moves up the food chain. Toxins may interact with each other in an additive, antagonistic, or synergistic way. The potency of a toxin is measure by the **lethal dose-50%** (**LD_{50}**), the concentration that is fatal to 50% of the test population. This is determined through a **dose-response curve**, which shows the percent of individuals affected as a function of concentration.

The **World Health Organization (WHO)** is an agency of the United Nations that collects and analyzes data on disease occurrence from member nations. WHO also coordinates public health programs and responses to international health emergencies. In the

United States, the **Centers for Disease Control and Prevention** monitors **notifiable diseases** and publishes weekly updates in the *Morbidity and Mortality Weekly Report.* Strategies for reducing environmental hazards include providing access to clean water, improving sanitation and hygiene, and limiting exposure to disease vectors. The public serves an important role by engaging in behaviors that limit disease spread or exposure to toxins and supporting policies that limit environmental hazards.

Attributions

Modified by Melissa Ha from the following sources:

- Disease and Epidemiology from *Microbiology* by OpenStax (licensed under CC-BY)
- Environmental Hazards & Human Health and Pests and Pesticides from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 5.4.7: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **15.7: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

CHAPTER OVERVIEW

6: Energy

Energy for lighting, heating and cooling our buildings, manufacturing products, and powering our transportation systems comes from a variety of natural sources. **Nonrenewable energy** is finite and cannot be replenished within a human timescale. Examples include nuclear energy and fossil fuels. In contrast **renewable resources** are replenished on shorter time scales, and it is thus possible to use them indefinitely. All energy sources have and some environmental and health cost, and the distribution of energy is not equally distributed among all nations.

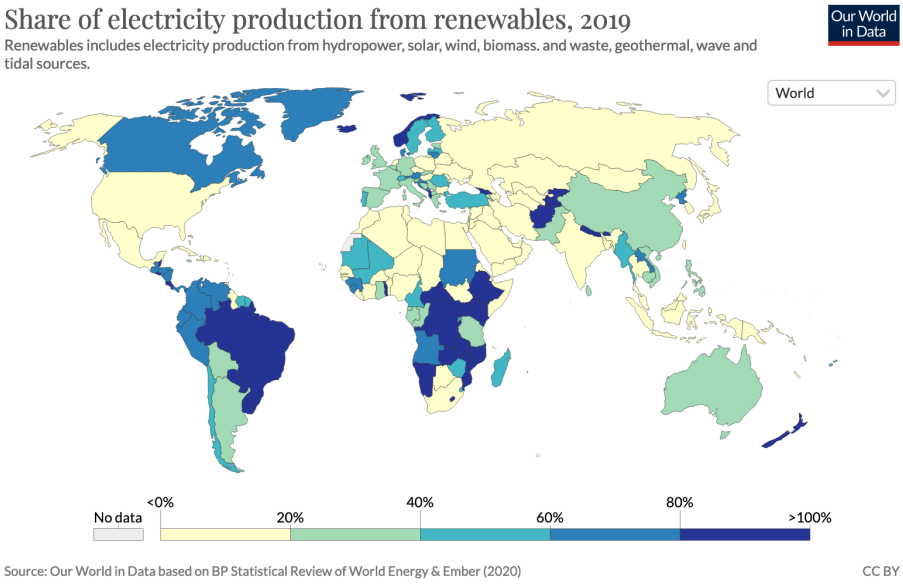


Figure 6.*a*: How much of the worlds energy comes from renewable sources. Graph by Our World in Data (CC-BY)

Attribution

Modified by Melissa Ha and Rachel Schleiger from Challenges and Impacts of Energy Use from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

6.1: Fossil Fuels
- 6.1.1: Types of Fossil Fuels and Formation
- 6.1.2: Mining, Processing, and Generating Electricity
- 6.1.3: Fossil Fuel Consumption
- 6.1.4: Consequences of Fossil Fuels
- 6.1.5: Data Dive- Global Fossil Fuel Consumption
- 6.1.6: Review

6.2: Nuclear Energy
- 6.2.1: Radioactive Isotopes
- 6.2.2: Generating Electricity with Nuclear Energy
- 6.2.3: Nuclear Energy Consumption
- 6.2.4: Consequences of Nuclear Energy
- 6.2.5: Data Dive- Global Nuclear Power Generation
- 6.2.6: Review

6.3: Renewable Energy

6.3.1: Renewable Energy History and Consumption
6.3.2: Wind Energy
6.3.3: Solar Energy
6.3.4: Geothermal Energy
6.3.5: Hydropower
6.3.6: Biofuels (Biomass Energy)
6.3.7: Energy Conservation
6.3.8: Data Dive- Global Renewable Energy Generation
6.3.9: Review

This page titled 6: Energy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

SECTION OVERVIEW

6.1: Fossil Fuels

Chapter Hook

Human use of coal dates back thousands of years, and archeologists have found evidence for its use and excavation across the globe. As such, it is impossible to know how coal's utility was discovered. In the Americas, Aztecs used coal for fuels and decoration. In Europe, Romans were noted as early as the late 2nd century AD to be extracting coal and using it to smelt iron ore and heat their famous public baths. Coals earliest reference in writings was by a Greek scientist Theophrastus (c.371-267 BC) about its usage in metalworking. In Asia, the Chinese are known to have been mining and using coal as far back as 3490 BC. Overall, there is a deep history of human usage of coal. As there is a limited supply, it regrettably won't last forever.



Figure 6.1.*a***:** Fossil charcoal in weathered coal. Image by James St. John in Wikimedia Commons (CC-BY2.0)

6.1.1: Types of Fossil Fuels and Formation

6.1.2: Mining, Processing, and Generating Electricity

6.1.3: Fossil Fuel Consumption

6.1.4: Consequences of Fossil Fuels

6.1.5: Data Dive- Global Fossil Fuel Consumption

6.1.6: Review

Attribution

Modified by Rachel Schleiger (CC-BY-NC).

This page titled 6.1: Fossil Fuels is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

6.1.1: Types of Fossil Fuels and Formation

Fossil fuels are nonrenewable sources of energy formed from the organic matter of plants and microorganisms that lived millions of years ago. This energy was originally captured via photosynthesis by living organisms such as plants, algae, and photosynthetic bacteria. Sometimes this is known as fossil solar energy since the energy of the sun in the past has been converted into the chemical energy within a fossil fuel. As discussed in Food Chains and Food Webs and Matter, the organic molecules store chemical energy, which is released when the higher energy (less stable) bonds in these molecules are broken to form lower energy (more stable) bonds. Fossil fuels are nonrenewable because their formation took millions of years. Furthermore, higher productivity in the ancient environment allowed for more fossil fuel accumulation, meaning that the fossil fuel reserves available now could not necessarily be regenerated millions of years in the future.

Fossil fuels are composed primarily of **hydrocarbons** (molecules of just carbon and hydrogen), but they contain lesser amounts nitrogen, sulfur, oxygen, and other elements as well. The precise chemical structures vary depending on the type of fossil fuel (coal, oil, or natural gas). The molecules in coal tend to be larger than those in oil and natural gas. Coal is thus solid at room temperature, oil is liquid, and natural gas is in a gaseous phase. Specifically, **coal** is a black or dark brown solid fossil fuel found as coal seams in rock layers formed from ancient swamp vegetation. Both oil and natural gas are fossil fuels found underground that formed from marine microorganisms. **Oil** (petroleum) is a liquid fossil fuel and consists of a variety of hydrocarbons while **natural gas** is a gaseous fossil fuel that consists of mostly methane and other small hydrocarbons.

Coal

Coal is the product of fossilized swamps, although some older coal deposits that predate terrestrial plants are presumed to come from algal buildups. Coal was formed when plant material is buried, heated, and compressed in oxygen-poor conditions over a long period of time (figure 6.1.1.*a*). Millions of years ago, continents were in different locations with different climates, and swamp-like vegetation covered many regions. When the vegetation died, it could not fully decompose due to oxygen-poor conditions. Instead, it formed peat (a brown substance high in organic content). The peat was buried and formed coal after millions of years of high pressure and temperature. The pressure was from the weight of sediments as well as from continental collisions.

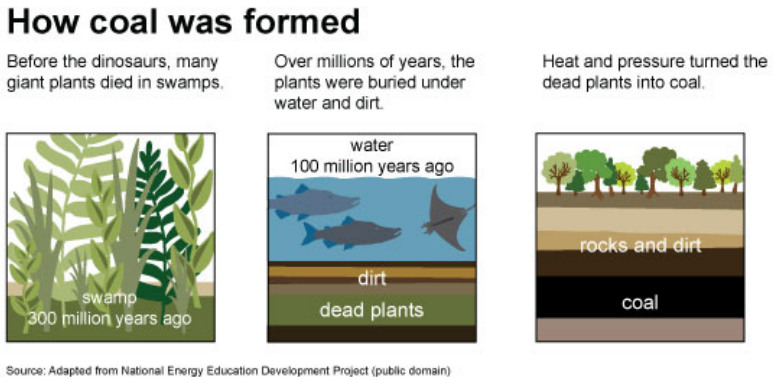


Figure 6.1.1.*a*: Coal was formed when large plants in swamps died 300 million years ago (before the dinosaurs). Over millions of years, this vegetation was buried under water and dirt (100 million years ago). Eventually, heat and pressure turned the dead plants into coal, which is found under layers of rock and dirt. Image from U.S. Energy Information Administration/National Energy Education Development Project (public domain).

There are several different types of coal ranging in quality (figure 6.1.1.*b*). The more heat and pressure that coal undergoes during formation, the greater is its fuel value and the more desirable is the coal. The general sequence of a swamp turning into the various stages of coal are as follows:

Swamp → Peat → Lignite → Subbituminous coal → Bituminous coal → Anthracitic coal → Graphite

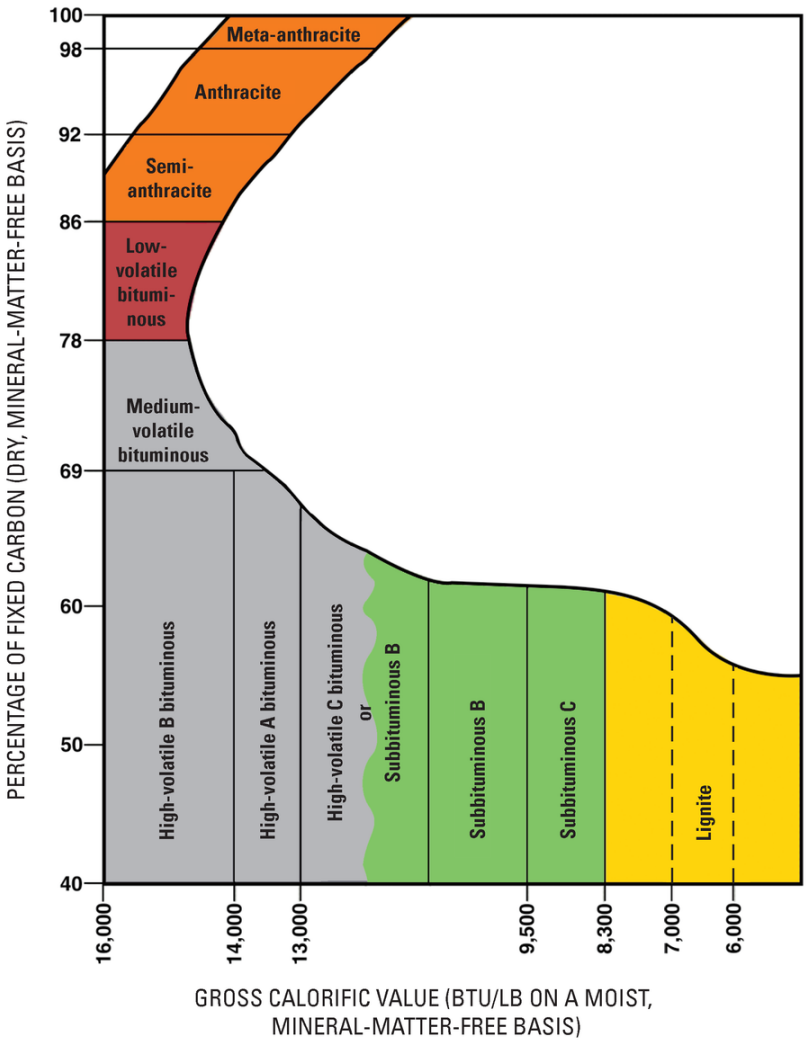


Figure 6.1.1.*b*: Coal rankings depend on energy content, measured as gross calorific value (how much energy is released from combustion) and carbon content that can be burned (percentage of fixed carbon). Anthracitic coal (orange) is the highest quality coal, with high energy and carbon content. Next in quality is bituminous coal (gray), subbituminous coal (green), and lignite (yellow). All three have less carbon content than anthracitic coal. Bituminous coal retains high energy content, but subbituminous coal and lignite have lower energy content. Image by USGS (public domain).

Specifically, peat compacts to form solid rock through a process called **lithification**, producing **lignite** (brown coal, a low-quality form of coal). With increasing heat and pressure, lignite turns to **subbituminous coal** and **bituminous coal**. Lignite, subbituminous coal, and bituminous coal are considered sedimentary rocks because they from from compacted sediments. At very high heat and pressure, bituminous coal is transformed to **anthracite**, a high-grade coal that is the most desirable coal since it provides the highest energy output (figure 6.1.1.*c*). Anthracite is considered a metamorphic rock because it has been compacted and transformed to the extent that it is denser than the other forms of coal and no longer contains sheet-like layers of sediments. With even more heat and pressure driving out all the components that evaporate easily and leaving pure carbon, anthracite can turn to **graphite**.



Figure 6.1.1.*c*: Anthracitic coal, the highest grade of coal.

Oil and Gas

Oil and natural gas formed from ancient marine microorganism (**plankton**). When plankton died, they were buried in sediments. As with coal, oxygen-poor conditions limited decomposition. As sediments continued to accumulate, the dead organisms were further buried. High temperature and pressure over millions of years ultimately produced oil and natural gas from these dead organisms.

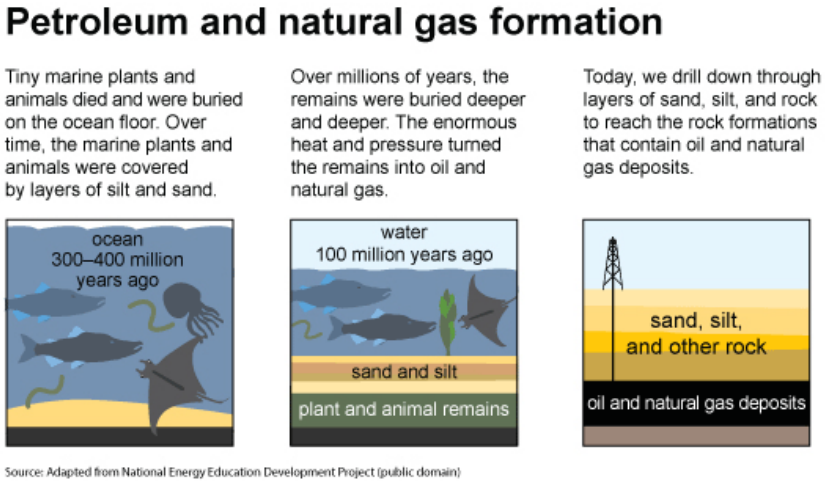


Figure 6.1.1.*d*: Petroleum (oil) and natural gas were formed from marine microorganisms. (The image text mentions tiny marine plants, but they were primarily algae and photosynthetic bacteria rather than plants.) These were covered by layers of silt and sand 300-400 million years ago. Over millions of years, the remains were buried deeper and deeper. They are pictured 100 million years ago. The enormous heat and pressure turned the remains into oil and natural gas. Now, oil and natural gas deposits are found underground and can be extracted via drilling through the layers of sand, silt, and rock. Image from U.S. Energy Information Administration/National Energy Education Development Project (public domain).

As the rock forms from the sediments that originally trapped the plankton, the oil and gas leak out of the **source rock** due to the increased pressure and temperature, and migrate to a different rock unit higher in the rock column. If the rock is porous and permeable rock, then that rock can act as a **reservoir** for the oil and gas. Petroleum is usually found one to two miles (1.6 – 3.2 km) below the Earth's surface, whether that is on land or ocean.

A **trap** is a combination of a subsurface geologic structure and an impervious layer that helps block the movement of oil and gas and concentrates it for later human extraction. Traps pool the fluid fossil fuels into a configuration in which extraction is more likely to be profitable, and such fossil fuels are called **conventional oil and natural gas** (figure 6.1.1.*e*). Extraction of oil or gas outside of a trap (**unconventional oil and natural gas**) is less efficient and more expensive; sometimes it is not economically viable at all (does not produce a profit). Examples of unconventional fossil fuels include oil shale, tight oil and gas, tar sands (oil sands), and coalbed methane.

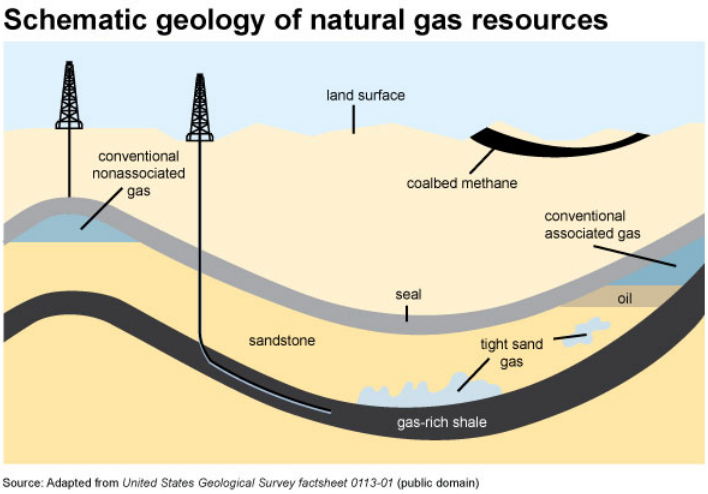


Figure 6.1.1.*e*: Conventional oil and natural gas deposits are trapped beneath impervious rock (gray). Conventional natural gas may be associated with oil or nonassociated. Coalbed methane and tight gas found in shale and sandstone are examples of unconventional fossil fuels. Image from USGS/EIA (public domain)

Oil Shale

Oil shale is a fine-grained sedimentary rock that sometimes contains **kerogen**, a solid material from which petroleum products can ultimately be manufactured. In order extract the fossil fuels, the material has to be mined and heated, which is expensive and typically has a negative impact on the environment.

Tight Oil and Natural Gas

Tight oil and natural gas are also trapped in shale rock, fine-grained sedimentary rocks with relatively high porosity and low permeability. They differ from oil shale in that they can be extracted through a process called hydraulic fracturing (fracking).

Similarly, fracking can be used to extract natural gas from tight sands, which are gas-bearing, fine-grained sandstones or carbonates (rocks made of minerals containing carbonate, CO_3^{2-}) with a low permeability.

Tar Sands

Tar sands, or oil sands, are sandstones that contain petroleum products that are highly viscous (like tar), and thus, can not be drilled and pumped out of the ground, unlike conventional oil (figure 6.1.1.*f*). The fossil fuel in question is bitumen, which can be pumped as a fluid only at very low rates of recovery and only when heated or mixed with solvents. Thus, injections of steam and solvent or mining of the tar sands for later processing can be used to extract the tar from the sands. (See related information about strip mining with respect to coal in Mining, Processing, and Generating Electricity.) Alberta, Canada is known to have the largest reserves of tar sands in the world.



Figure 6.1.1.*f*: Tar sandstone from the Miocene Monterrey Formation of California.

Coalbed Methane

Some natural gas is also found associated with coal deposits (**coalbed methane**), consisting of methane produced during coal formation.

Attributions

Modified by Melissa Ha from the following sources:

- Challenges and Impacts of Energy Use and Non-Renewable Energy Sources from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Fossil Fuels from *An Introduction to Geology* by Chris Johnson et al. (licensed under CC-BY-NC-SA)
- Shale Gas 101. Office of Fossil Energy. U.S. Department of Energy. Accessed 01-12-2021. (public domain)
- The Process of Unconventional Natural Gas Production. 2021. United States Environmental Protection Agency. Accessed 01-12-2021. (public domain)

This page titled 6.1.1: Types of Fossil Fuels and Formation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **16.1: Types of Fossil Fuels and Formation** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

6.1.2: Mining, Processing, and Generating Electricity

Fossil fuels must be extracted or mined before use, and the specific method depends on the type of fossil fuel. Coal and natural gas and primarily used for electricity generation, while petroleum is refined to produce fuel for vehicles, planes, and heating as well as other products.

Coal

Mining

Coal is extracted by two principal methods, of which there are many variants: surface mining or subsurface mining. **Surface mining** uses large machines to remove the soil and layers of rock known as **overburden** to expose coal seams that are close to the Earth's surface (figure 6.1.2.a). **Strip mining** is a type of surface mining in which overburden is sequentially removed from each stretch (strip) of land. Once overburden is removed from the first strip, coal is removed. Overburden from the second strip is then deposited into the first strip, and coal is removed from the second strip. Overburden from the third strip is then placed in the first strip, and so on. **Mountaintop removal** is a more destructive type of surface mining in which all of the overburden is removed with explosives, revealing the entire coal seam at once (figure 6.1.2.b). The large mass of overburden (the mountaintop) is dumped into a nearby valley, and the coal is them removed.



Figure 6.1.2.a: Surface mining of coal in Wyoming. Image by Bureau of Land Management (public domain).

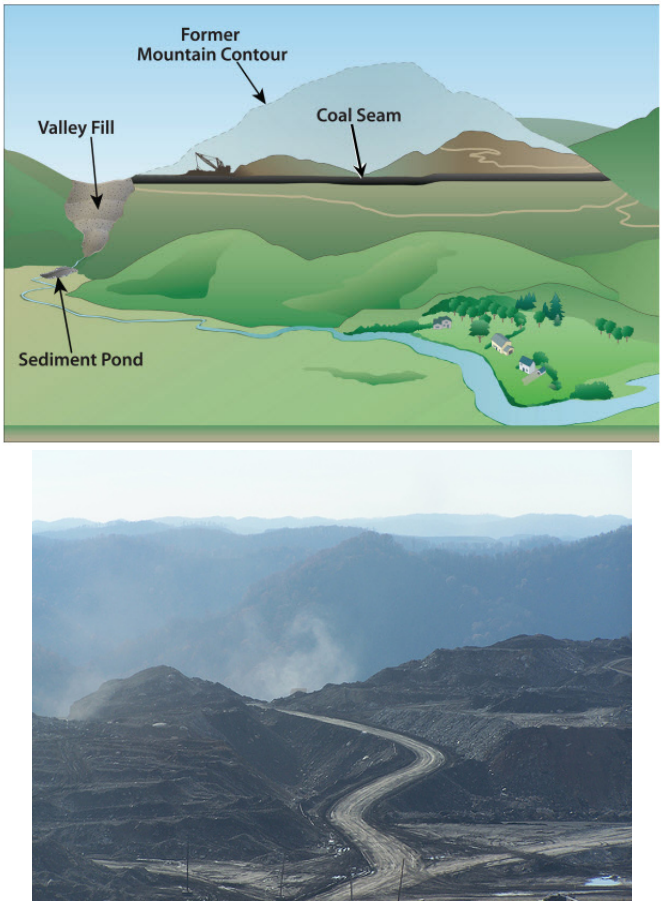


Figure 6.1.2.*b*: Left: Mountaintop removal uses explosives to loosen overburden at the top of a mountain, which is then displaced into a nearby valley (valley fill). The underlying coal seam is then extracted. A sediment pond collects soil that erodes from the valley fill. Image by EPA (public domain). Right: A mountaintop removal site. Image by JW Randolph (public domain)

Subsurface mining (deep mining) employs underground tunnels to access deeper deposits (figure 6.1.2.*c*). Some underground mines are thousands of feet deep, and extend for miles. Miners ride elevators down deep mine shafts and travel on small trains in long tunnels to get to the coal. The miners use large machines that dig out the coal. In **drift mines**, a tunnel is dug horizontally into the side of a mountain. In **slope mines**, this tunnel is diagonal. In **shaft mines**, elevators are used to move coal through vertical tunnels.

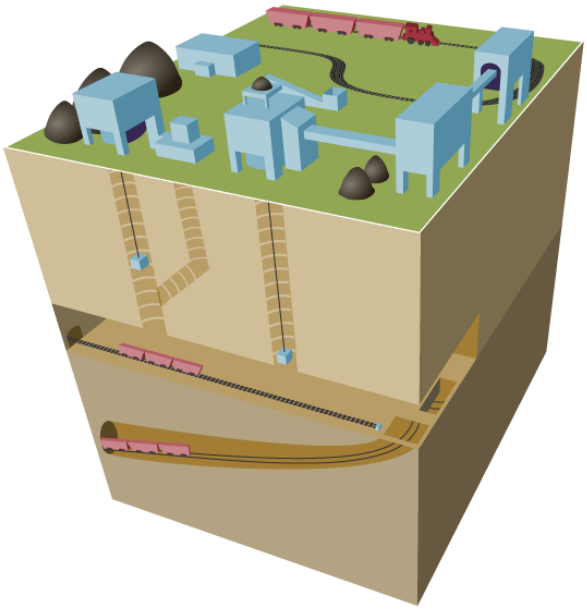


Figure 6.1.2.*c*: Subsurface mining involves digging tunnels to access coal deposits that are deep underground. Image by National Energy Education Development Project/EIA (public domain)

Processing of Coal

Once mined, coal may go to a preparation plant located near the mining site where it is cleaned and processed to remove impurities such as rocks and dirt, ash, sulfur, and other unwanted materials. This process increases the amount of energy that can be obtained from a unit of coal, known as its heating value.

Transport of Coal

Finally, the mined and processed coal must be transported. Transportation can be more expensive than mining the coal. Nearly 70% of coal delivered in the United States is transported, for at least part of its trip, by train (figure 6.1.2.*e*). Coal can also be transported by barge, ship, or truck. Coal can also be crushed, mixed with water, and sent through a slurry pipeline. Sometimes, coal-fired electric power plants are built near coal mines to lower transportation costs.



Figure 6.1.2.*e*: A freight train loaded with coal briquettes, formed from compressed coal dust, in Morwell, Victoria, Australia. Image and caption (modified) by CSIRO (CC-BY).

Generating Electricity from Coal

Once at the power plant, coal is first pulverized into a fine powder and then mixed with hot air and blown into a furnace (figure 6.1.2.f). This allows for the most complete combustion (burning) and maximum heat release. Purified water, pumped through pipes inside a boiler, is turned into steam by the heat from the combustion of coal. The high pressure of the steam pushing against a series of giant turbine blades turns the turbine shaft. The turbine shaft is connected to the shaft of the generator, where magnets spin within wire coils to produce electricity. After doing its work in the turbine, the steam is drawn into a condenser, a large chamber in the basement of the power plant. In this important step, millions of gallons of cool water from a nearby source (such as a river or lake) are pumped through a network of tubes running through the condenser. The cool water in the tubes converts the steam back into water that can be used over and over again in the plant. The cooling water is returned to its source without any contamination except at a higher temperature than when first extracted from the river or lake.

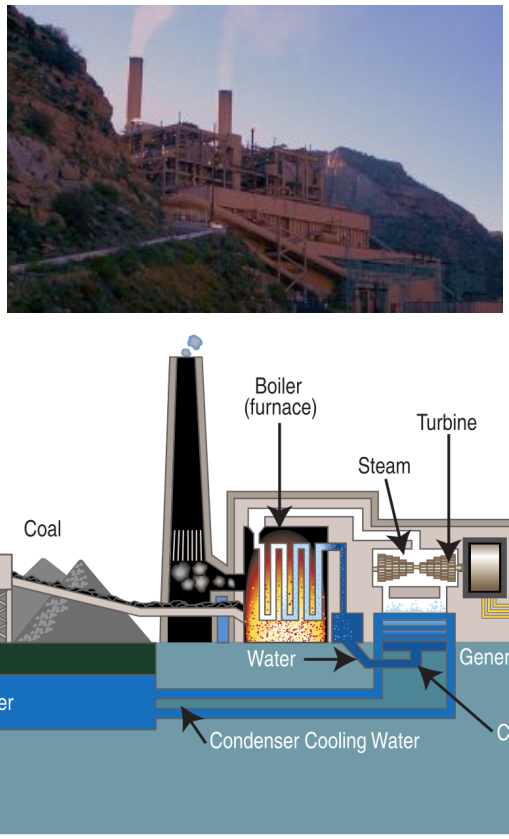
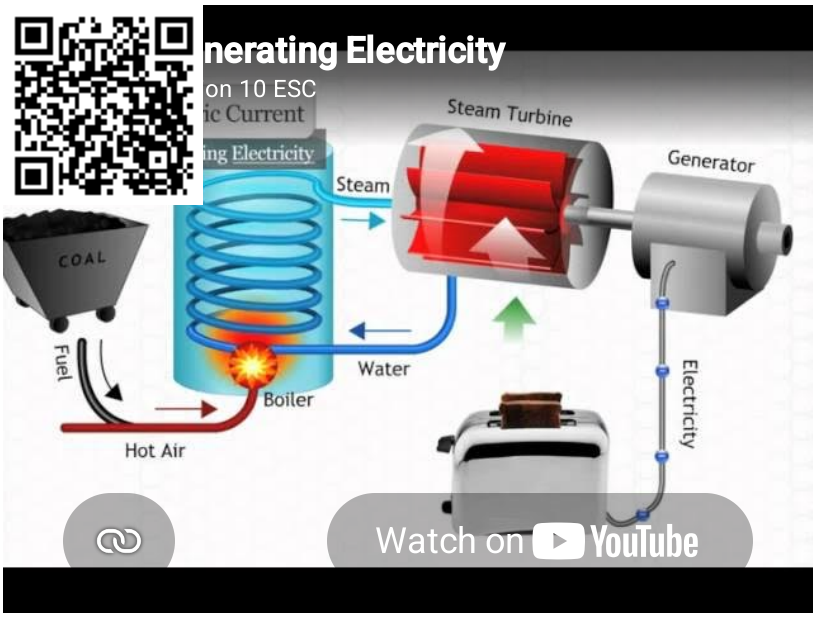


Figure 6.1.2.f: Left: Coal power plant in Helper, Utah. Right: Diagram of a typical steam-cycle coal power plant (proceeding from left to right). Coal enters the boiler (furnace), where it is burned. This boils water, producing steam. The steam turns a turbine, powering a generator. The resultant electricity is passed through a transformer (which changes the voltage) and is then sent through transmission lines. The steam cools and condenses back to liquid water in the condenser. This is facilitator by cooling water from a nearby river. Right image by Tennessee Valley Authority (public domain).

This video shows how heat energy can be used to generate electricity.



Oil and Natural Gas

Extraction of Conventional Oil and Natural Gas

Conventional oil and natural gas are contained under a trap (cap rock). Because natural gas consists of lighter molecules that are in a gaseous form at moderate temperatures, it is found on top of the oil, which may be floating on groundwater. To access conventional oil and natural gas, the trap is first pierced. Initially, they are under high enough pressure, and this drives them out of the well (**primary recovery**). Next, water (or gas) is injected to force more fossil fuels out (**secondary recovery**). Finally, enhanced oil recovery (**tertiary recovery**) may be used to extract further oil by applying heat (injecting steam) or injecting carbon dioxide, other gases, or larger molecules. For example, carbon dioxide causes the oil to thin and expand, making it easier to remove from the rocks. Note that secondary recovery simply increases pressure inside the reservoir whereas tertiary recovery changes the properties of the oil, making it easier to extract (figure 6.1.2.*g*). Each stage of recovery is increasingly expensive, and extraction from a well continues as long as it remains profitable.

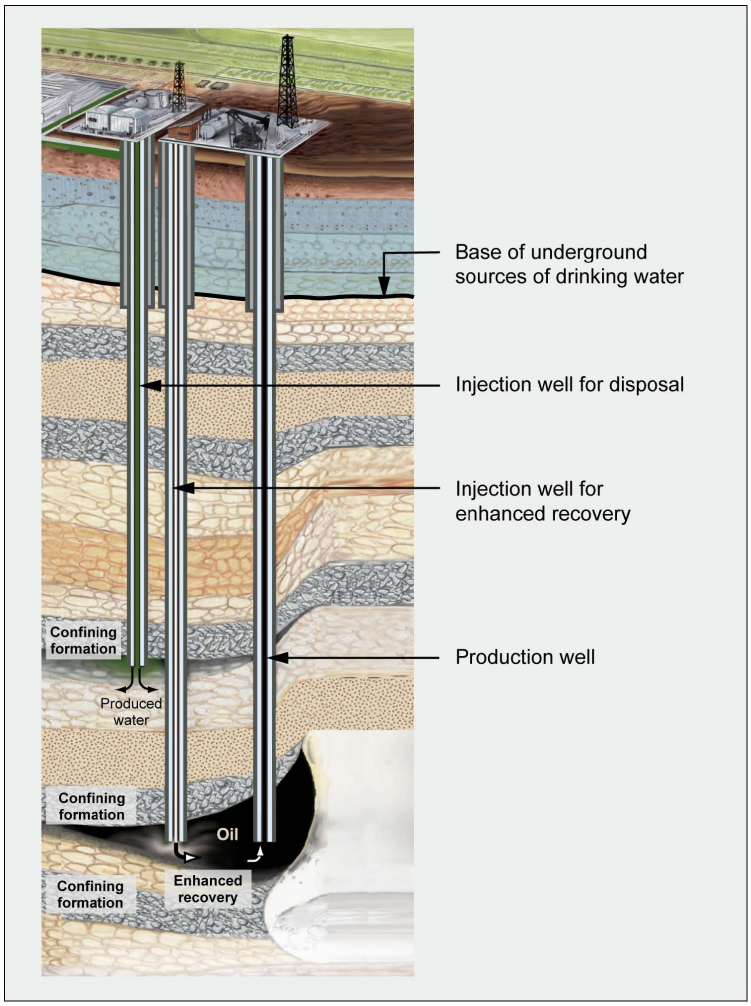


Figure 6.1.2.*g*: Injection wells transfer water, carbon dioxide, or other substances to an oil deposit increase pressure or change the oil's properties, facilitating extraction. On the right is the production well, through which extracted oil flows. This is facilitated by an injection well for enhanced recovery. On the left is an injection well for disposal through with waste water (produced water) is stored underground. Several confining formations trap substances underground. Closer to the top is the base of underground sources of drinking water, meaning that all drinking water is extracted from above this point. Image by Government Accountability Office (public domain).

Oil is mainly obtained by drilling either on land (onshore) or in the ocean (offshore). Early offshore drilling was generally limited to areas where the water was less than 300 feet deep. Oil and natural gas drilling rigs now operate in water as deep as two miles. Floating platforms are used for drilling in deeper waters (figure 6.1.2.*h*). These self-propelled vessels are attached to the ocean floor using large cables and anchors. Wells are drilled from these platforms which are also used to lower production equipment to the ocean floor. Some drilling platforms stand on stilt-like legs that are embedded in the ocean floor. These platforms hold all required drilling equipment as well as housing and storage areas for the work crews. Offshore production is much more expensive than land-based production.



Figure 6.1.2.*h*: A platform for offshore oil drilling. Image by Pixabay/keridjackson (Pixabay license).

Extraction of Unconventional Oil and Natural Gas

Tight oil and natural gas trapped in shale as well as natural gas in tight sands are extracted via **hydraulic fracturing**, informally referred to as “fracking”. This process uses explosives to create new fractures in these low-permeability rocks as well as increase the size, extent, and connectivity of existing fractures and then applies high-pressure fluid. First, a drill permeates the rock layers and then proceeds horizontally. Explosives then fracture rocks, freeing oil and natural gas. Finally, water, sand, and chemicals and injected, which flush out oil and natural gas (figure 6.1.2.*i*).

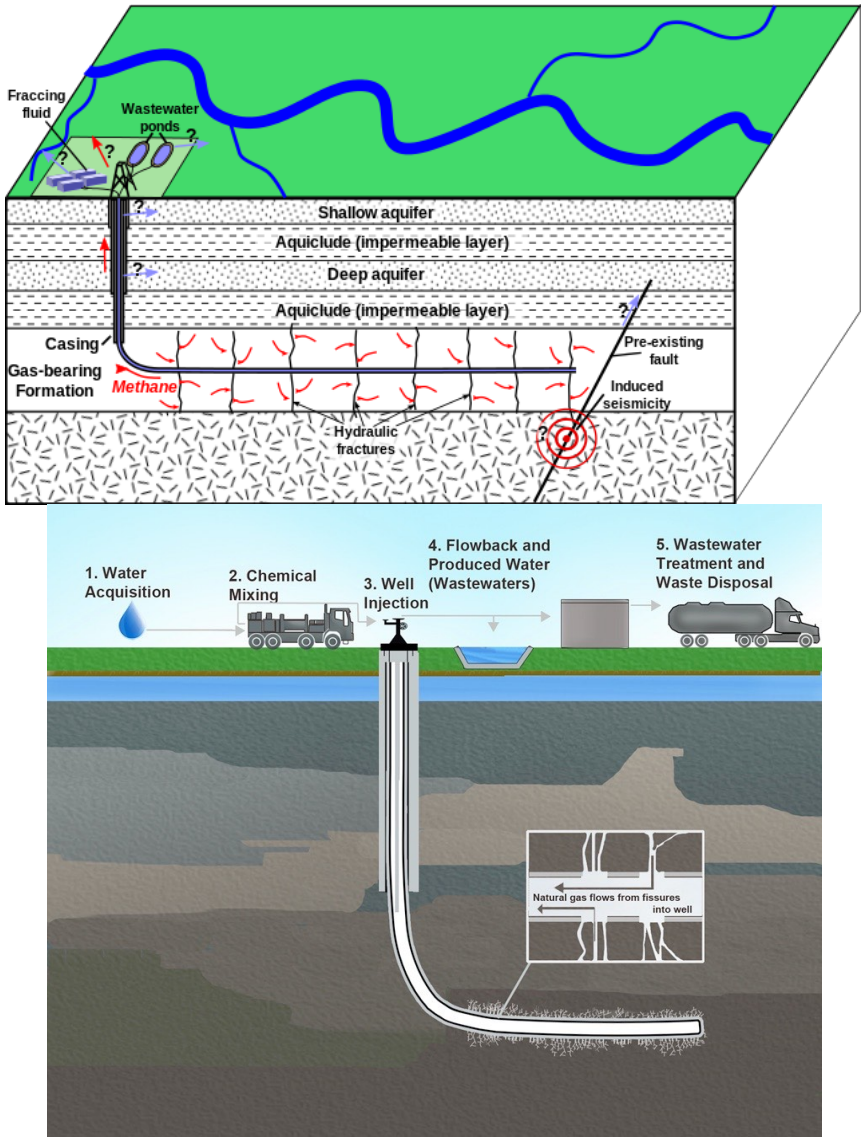


Figure 6.1.2.*i*: Two diagrams of fracking. Water is mixed with sand and chemicals and then injected into shale trapping tight oil and natural gas or tight sand. This flushes fossil fuels from the fissures that were previously created by explosives. Top: This section shows four underground layers. From top to bottom, they are the shallow aquifer, aquiclude (impermeable layer), deep aquifer, and another aquiclude. Fracking fluid is injected underground through a well surrounded by casing into the gas-bearing formation. Methane (red arrows) might escape from hydraulic fractures in this formation. Additionally, when the fractures intersect with a pre-existing fault, induced seismicity (earthquakes) is possible. Aboveground, fracking fluid is stored in wastewater ponds. Blue arrows and question marks indicate places where toxic wastewater may escape and contaminate groundwater such as from storage containers, wastewater ponds, casing, or faults. Bottom: Proper treatment and disposal of the wastewater is needed to limit environmental impact. 1: Water is acquired. 2: Chemicals are mixed. 3: Fracking fluid is injected into the well. Here, natural gas flows from fissures into the well. 4: Fracking results in flowback and produced water (wastewater). 5: Wastewater undergoes treatment and disposal. Bottom image by USGS (public domain).

As mentioned previously, bitumen in tar sands can be extracted by injecting steam, or they may be mined for later processing. Tar sands can mined through strip mining or **open-pit mining**, a type of surface mining that involves forming a progressively deeper hole. The walls of the pit are as steep as can safely be managed. A steep wall means there is less waste overburden to remove and is an engineering balance between efficient mining and mass wasting. Oil shale is extracted by strip mining, creating subsurface mines, or open-pit mining. Oil shale can be burned directly like coal or baked in the presence of hydrogen to extract liquid petroleum (figure 6.1.2.*j*).



Figure 6.1.2.j: Underground mining of oil shale in Estonia.

Refining Crude Oil

The result of oil recovery is crude oil (petroleum), which contains many types of hydrocarbons as well as some unwanted substances such as sulfur, nitrogen, oxygen, dissolved metals, and water all mixed together. Unprocessed crude oil is therefore, not generally useful in industrial applications and must first be separated into different useable products (**petrochemicals**) at a refinery. Gasoline (petrol), diesel, tar, and asphalt are examples of petrochemicals.

Fractional distillation is the key process used in oil refineries to separate the components of crude oil. During fractional distillation, crude oil is heated and then allowed to cool. The heaviest compounds sink to the bottom as residues. Components of vaporized crude oil condense at different levels in the distillation column depending on their boiling points, which is primarily due to their molecular sizes. The heaviest compounds (condense near the bottom of the column, where the temperature is still high. Lighter compounds condense at cooler temperatures higher up in the column. Some compounds remain as gases at the top of the column (figure 6.1.2.k).

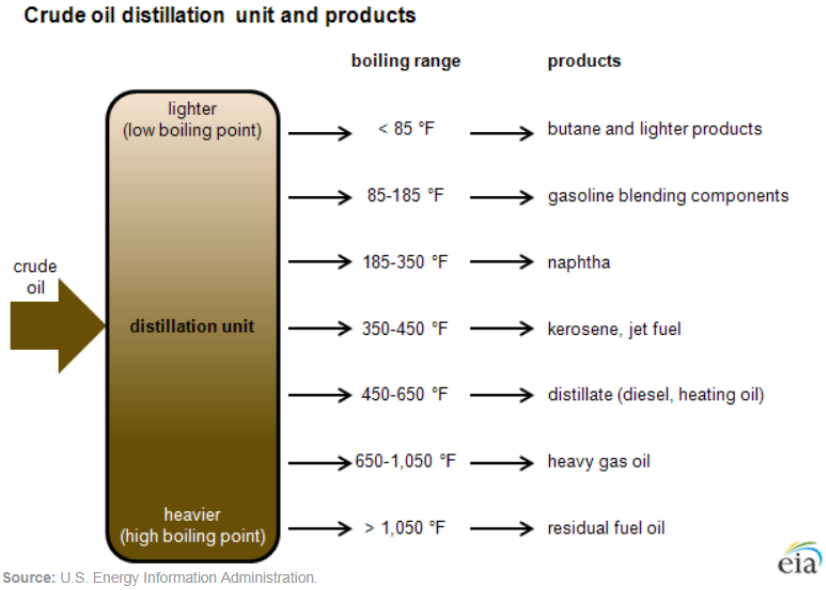


Figure 6.1.2.k: The process of fractional distillation involves heating crude oil and allowing the components to cool. As they move up the distillation column, they condense at different levels based on their boiling points. The heaviest compounds have the highest boiling points and condense at the bottom of the column while lighter compounds (low boiling point) condense at the top. From lowest to highest boiling point the petrochemicals produced are residual fuel oil (1050 °F), heavy gas oil (650-1050 °F), diesel and heating oil (450-650 °F), kerosene and jet fuel (350-450 °F), naphtha (used to make gasoline, solvents, cleaning solutions, etc.; 185-350 °F), gasoline blending components (85-185 °F), and butane and lighter products (< 85 °F). Image by U.S. Energy Information Administration (public domain).

The video below explains the process of fractional distillation. The labeled distillation column at 3:00 shows heated crude oil (400 °C) separating into various petrochemicals. From bottom to top, they are bitumen (> 350 °C), diesel (250-350 °C), kerosene

(160-250 °C), naphtha (70-160 °C), petrol (20-70 °C), and gas (< 20 °C).



Conversion is the chemical processing in which some of the fractions (produced from fractional distillation) are transformed in to other products. For example, a refinery can turn diesel fuel into gasoline depending on the demand for gasoline. Conversion can involve breaking larger hydrocarbon chains into smaller ones (cracking), combining smaller chains into larger ones (unification), or rearranging the molecules to created desired products (alteration).

Treatment is done to the fractions to remove impurities such as sulfur, nitrogen and water among others. Refineries also combine the various fractions (processed and unprocessed) into mixtures to make desired products. For example, different mixtures of hydrocarbon chains can create gasolines with different octane ratings, with and without additives, lubricating oils of various weights and grades (WD-40, 10W-40, 5W-30, etc.), heating oil, and many others. The products are stored onsite until they can be delivered to various markets such as gas stations, airports and chemical plants.

A 42 U.S. gallon barrel of crude oil yields about 45 gallons of petroleum products because of refinery processing gain (figure 6.1.2.*l*). This increase in volume is similar to what happens to popcorn when it is popped. Gasoline makes up the largest fraction of all petroleum products obtained. Other products include diesel fuel and heating oil, jet fuel, petrochemical feedstocks (to manufacture plastics, synthetic rubber, or other chemicals), waxes, lubricating oils, and asphalt.

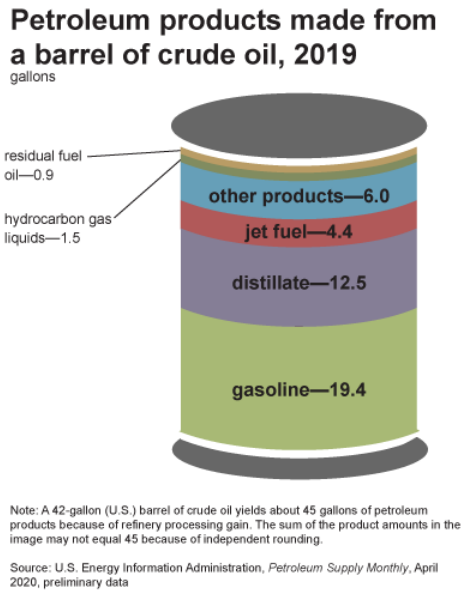


Figure 6.1.2.*l*: Main products (measured in gallons) made from a barrel of crude oil in 2019. These include residual fuel oil (0.9 gallons), hydrocarbon gas (1.5), other products (6.0), jet fuel (4.4), distillate (12.5), and gasoline (19.4). Note: A 42-gallon (U.S.) barrel of crude oil yields about 45 gallons of petroleum products because of refinery processing gain. The sum of the product amounts in the image may not equal 45 because of independent rounding. Image by EIA (public domain).

Transporting Oil and Natural Gas

After the refinery, the gasoline and other fuels created are ready to be distributed for use. A system of pipelines runs throughout the United States to transport oil and fuels from one location to another. There are pipelines that transport crude oil from the oil well to the refinery. At the refinery, there are additional pipelines that transport the finished product to various storage terminals where it can then be loaded onto trucks for delivery, such as to a gas station.

Once natural gas is produced from underground rock formations, it is sent by pipelines to storage facilities and then on to the end user. The United States has a vast pipeline network that transports gas to and from nearly any location in the lower 48 states. There are more than 210 natural gas pipeline systems, using more than 300,000 miles of interstate and intrastate transmission pipelines (figure 6.1.2.m). Compressor stations that maintain pressure on the natural gas to keep it moving through the system. There are more than 400 underground natural gas storage facilities that can hold the gas until it is needed back in the system for delivery.



Figure 6.1.2.m: Construction of the controversial Dakota Access Pipeline, which stretches from North Dakota to Illinois. The pictured portion of the pipeline is in Central Iowa. You can learn more about the 2016 protests against this pipeline here. Image by Dakota Access Pipe Line (CC-BY).

Generating Electricity from Oil or Natural Gas

Natural gas is burned to produce electricity following the same general process used in a coal power plant (figure 6.1.2.n). Oil is occasionally used to generate electricity as well.

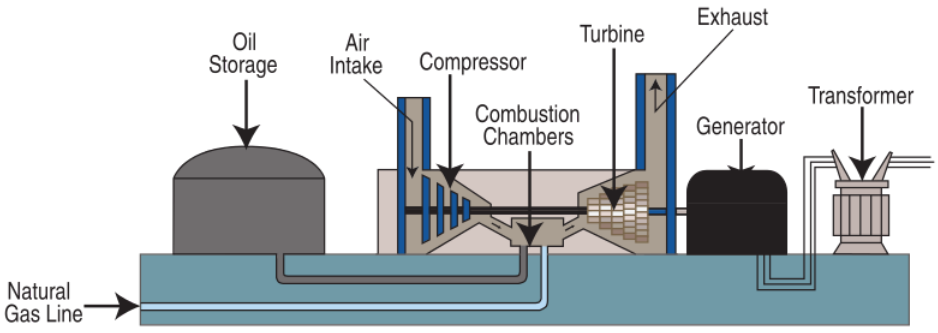


Figure 6.1.2.n: This combustion chamber burns either natural gas or oil. Fuel flows through a natural gas line or from oil storage into the combustion chamber. Air passes through the air intake and is compressed in the compressor. Natural gas and compressed air are mixed with compressed air in the combustion chamber and burned. High-pressure combustion gases spin the turbine, which drives the generator. The resultant electric current is then passed through a transformer, which alters voltage. Image and caption (modified) by Tennessee Valley Authority (public domain).

Attributions

Modified by Melissa Ha from the following sources:

- Non-Renewable Energy Sources from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Fossil Fuels and Mining from *An Introduction to Geology* by Chris Johnson et al. (licensed under CC-BY-NC-SA)
- Chapter 4: Non-Renewable Energy from *Introduction to Environmental Science: 2nd Edition* (2018) Biological Sciences Open Textbooks by Zehnder, Caralyn; Manoylov, Kalina; Mutiti, Samuel; Mutiti, Christine; VandeVoort, Allison; and Bennett, Donna (licensed under CC-BY-NC-SA).
- Fossil Energy Study Guide: Natural Gas. 2014. U.S. Department of Energy. Accessed 01-12-2021.
- Fossil Energy Study Guide: Oil. 2013. U.S. Department of Energy. Accessed 01-12-2021.

This page titled 6.1.2: Mining, Processing, and Generating Electricity is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **16.2: Mining, Processing, and Generating Electricity** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

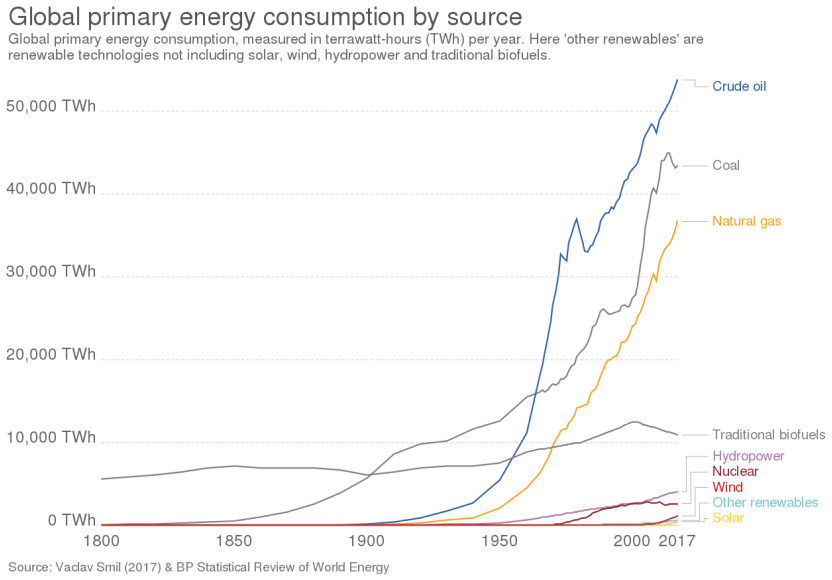
6.1.3: Fossil Fuel Consumption

The U.S. and the world overall heavily depend on fossil fuels. In 2019, fossil fuels contributed to 62.6% of electricity generation in the U.S. with coal contributing 23.4% and natural gas contributing 38.4% (table 6.1.3.a). Note that oil (petroleum) is primary used for transportation and thus only contributes a fraction of a percentage to electricity generation. With respect to total energy consumption, the world continues to rely on crude oil more than any other energy source (33.1%) followed by coal (27%) and natural gas (24.3%; figure6.1.3.a).

Table 6.1.3.a**:** Contribution of Each Energy Source to Electricity Generation in the U.S. in 2019

Energy source	Contribution to Electricity Generation (%)
Fossil fuels (total)	**62.6%**
Natural gas	38.4%
Coal	23.4%
Petroleum (total)	0.4%
Other gases	0.3%
Nuclear	**19.6%**
Renewables (total)	**17.6%**
Hydropower	7.0%
Wind	7.1%
Biomass (total)	1.4%
Solar (total)	1.7%
Geothermal	0.4%

Table modified from U.S. Energy Information Administration (public domain).



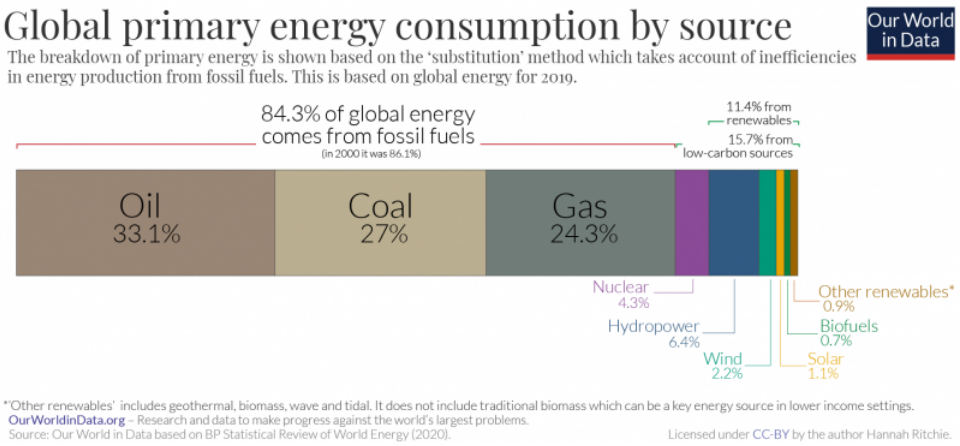


Figure6.1.3.a**:** The top graph shows global energy consumption by source shows energy consumption in terrawatt-hours (TWh) on the y-axis and time in years (1800-2017) on the X-axis. The bottom graph shows the percentage that each energy source contributes to global energy consumption in 2019. Global consumption of most energy sources has increased over time. Consumption of crude oil is the highest (33.1% of global energy consumption in 2019), and it increased rapidly in the 1950s, surpassing coal consumption in the early 1960s. Coal is now the second most-consumed energy source (27%), but its global consumption has declined in recent years. Natural gas is the third most-consumed energy source (24.3%) and has also been increasing. Coal and natural gas consumption surpassed that of traditional biofuels in the early 1900s and 1970, respectively. Nuclear energy consumption has decreased recently, accounting for 4.3% of global energy consumption in 2019. Consumption of renewable energy overall (11.4%) is low compared to fossil fuels (84.3% combined), but it has generally increased in recent years. Images by Our World in Data (CC-BY).

Coal has been used by humans for at least 6000 years, mainly as a fuel source. Coal resources in Wales are often cited as a primary reason for the rise of Britain (and later, the United States) in the Industrial Revolution. Coal electricity traces its origins to the early 20th Century, when it was the natural fuel for steam engines given its abundance, high energy density and low cost. Coal is the largest domestically produced source of energy. At the end of 2018, BP estimated at 734,903 million tonnes, with nearly 23.7% of that in the United States. It is a major fuel resource that the United States controls domestically.

Proven oil reserves (or natural gas reserves) refers to the amount of oil or natural gas that can be extracted economically with current methods (such as conventional wells or fracking). The U.S. Energy Administration estimates that there are enough liquid fuels to last through 2050 (and they include biofuels in this projection). In 2016, BP projected that proven reserves of oil and natural gas can support global demands for another 50 years. Additionally, they estimate that coal reserves can last another 115 years.

Scientists and policy-makers often discuss the question of when the world will reach **peak oil production**, the point at which oil production is at its greatest and then declines. It was initially predicted that global peak oil would be reached in 2000, but oil production and consumption continue to rise. The United States, however, already passed peak oil production in 1970.

The concentration of oil reserves in a few regions of the world makes much of the world dependent on imported energy for transportation (figure6.1.3.b). The rise in the price of oil in the last decade makes dependence on imported energy for transportation an economic as well as an energy issue. The United States spent $304.9 billion on oil imports in 2019. The United States has become more and more dependent on foreign oil since 1970 when our own oil production peaked.

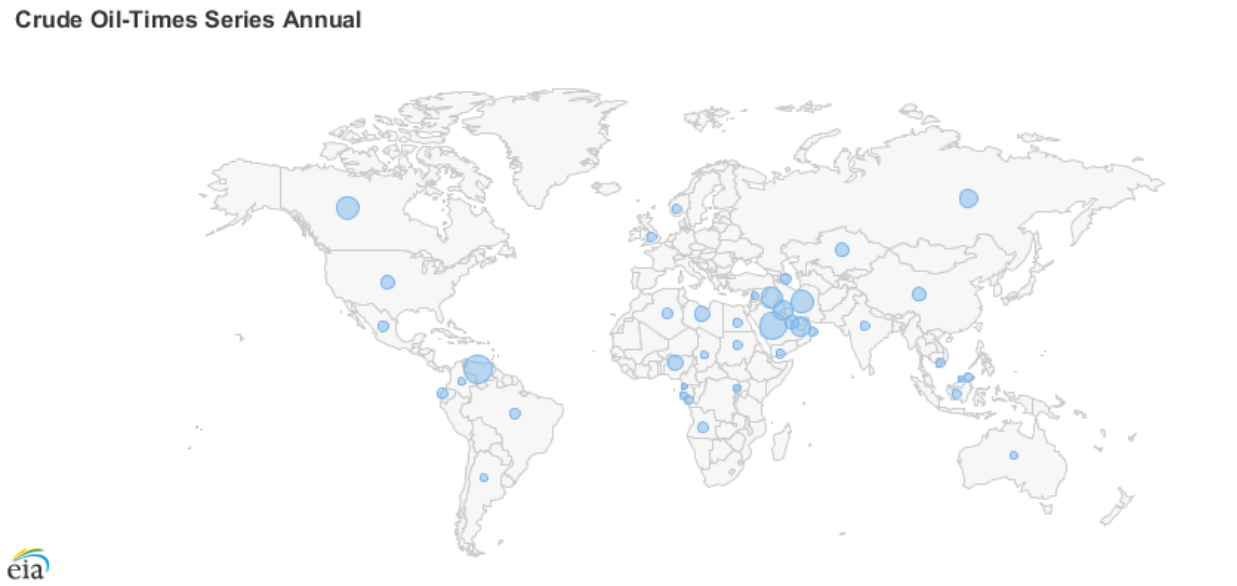


Figure6.1.3.*b*: Global reserves of oil (top) and natural gas (bottom) by country in 2017 and 2020, respectively. The blue circle in each country is scaled to the size of the reserve. Both fossil fuels are abundant in the Middle East. Russia also has a rich supply of natural gas. Images from EIA (public domain).

The major holder of oil reserves is the **Organization of Petroleum Exporting Countries,** (OPEC). As of 2024, there were 12 member countries in OPEC: Algeria, Congo, Equatorial Guinea, Gabon, Iran, Iraq, Kuwait, Libya, Nigeria, Saudi Arabia, the United Arab Emirates, and Venezuela. OPEC attempts to influence the amount of oil available to the world by assigning a production quota to each member except Iraq, for which no quota is presently set.

Overall compliance with these quotas is mixed since the individual countries make the actual production decisions. All of these countries have a national oil company but also allow international oil companies to operate within their borders. They can restrict the amounts of production by those oil companies. Therefore, the OPEC countries have a large influence on how much of world demand is met by OPEC and non-OPEC supply.

This pressure has lead the United States to developing policies that would reduce reliance on foreign oil such as developing additional domestic sources and obtaining it from non-Middle Eastern countries such as Canada, Mexico, Venezuela, and Nigeria.

However, since fossil fuel reserves create jobs and provide dividends to investors, a lot is at stake in a nation that has oil reserves. Oil wealth may be shared with the country's inhabitants or retained by the oil companies and dictatorships, such as in Nigeria prior to the 1990s.

Attribution

Modified by Melissa Ha from the following sources:

- Challenges and Impacts of Energy Use and Non-Renewable Energy Sources from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Hannah Ritchie (2017) - "Fossil Fuels". Published online at OurWorldInData.org. Accessed 01-16-2021. (licensed under CC-BY)

This page titled 6.1.3: Fossil Fuel Consumption is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **16.3: Fossil Fuel Consumption** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

6.1.4: Consequences of Fossil Fuels

Benefits of Using Fossil Fuels

The world is heavily dependent on fossil fuels, and existing infrastructure and technologies facilitate their continued use. An advantage of using coal for electricity is that it plentiful and inexpensive, especially in the United States, which has larger coal reserves than any other country. Furthermore, coal mining is a source of jobs and tax income. Coal's economic advantage is dwindling, however, as technologies associated with renewable sources of energy, such as solar and wind, become more efficient and inexpensive. The U.S. Energy Information Administration compared the levelized cost of electricity (LCOE) for technologies that will begin use in 2023, and the cost of coal-generated electricity exceeded that of many renewable sources (figure 6.1.4.*a*). The LCOE accounts for the building and operation costs of power plants, solar panels, wind turbines, etc.

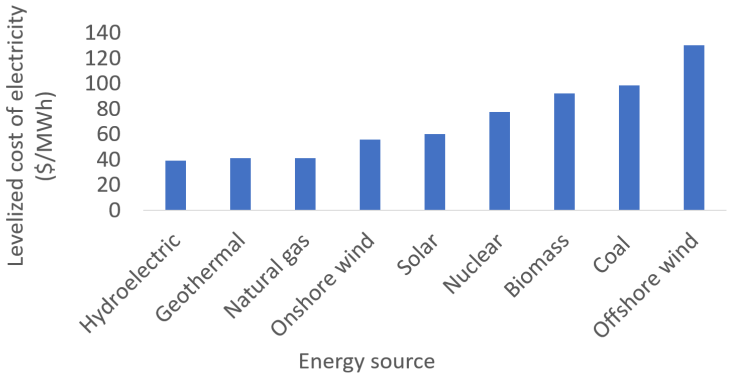


Figure 6.1.4.*a*: The **levelized cost of electricity** (**LCOE**) for various energy sources. Of these, hydroelectric, geothermal, natural gas, onshore/offshore wind, solar, and biomass are considered renewable. Nuclear and coal are nonrenewable energy sources. Values are in 2018 dollars per **megawatt hour** ($/MWh). As a reference, the average U.S. household, uses about 0.909 MWh of electricity per month. These values are for the newest technologies that will being use in 2023. Hydroelectric energy is the least expensive, at about $39/MWh. Hydroelectric energy is the least expensive, at about $39/MWh. Next is geothermal ($41/MWh), natural gas ($41/MWh), onshore wind ($56/MWh), solar ($60/MWh), nuclear ($78/MWh), biomass ($92/MWh), coal ($99/MWh), and offshore wind ($130/MWh). Graph by Melissa Ha (CC-BY-NC) using data U.S. Energy Information Administration, Annual Energy Outlook 2019 (public domain).

Oil and natural gas continue to meet global energy needs. Despite expansions in renewable energy use, no alternative energy source is currently sufficient to replace oil and natural gas. (A combination of different renewable sources could be possible in the future.) While the United States does rely on imported oil, it continues to produce some oil and natural gas (mostly through fracking), bolstering U.S. energy independence. Local and state economics in regions rich in oil and natural reserves depend on continued extraction of these fossil fuels.

While all fossil fuels harm cause some degree of environmental harm, natural gas is a preferred fossil fuel for electricity generation when considering its environmental impacts. When burned, coal emits nearly double the carbon dioxide that natural gas does. Additionally, much less nitrogen oxides and sulfur dioxide (both air pollutants) are emitted from burning natural gas. It also does not produce ash as coal does (see below).

Health and Environmental Impacts

The negative impacts of fossil fuel use begin with the extraction of the resource. Fossils fuels are often located far from where they are utilized so they need to be transported by pipeline, tankers, rail or trucks. These all present the potential for accidents, leakage, and spills. Additional negative impacts are associated with processing, electricity generation, and disposal of the waste generated.

Coal Mining and Usage

Surface mining of coal disrupts local ecosystems above coal deposits as overburden is removed to access them (figure 6.1.4.*b*). In mountaintop removal, a large volume of overburden is dumped over nearby habitats, causing further destruction (figure 6.1.4.*c*). Mountaintop removal has affected large areas of the Appalachian Mountains in West Virginia and Kentucky. Habitat loss from coal mining decreases biodiversity, resulting in a loss of ecosystem services. Both surface and subsurface mining expose rocks that can

contain contaminants, such as heavy metals or sulfates, which them leach into streams or other bodies of water. This not only harms aquatic life, but it also disrupts nutrient cycling. One of the largest environmental impacts of subsurface mining may be the methane gas that must be vented out of mines to make the mines a safe place to work. Methane gas is a potent greenhouse gas and contributes to climate change. Finally, the process of mining ultimately compacts soil. This combined with the loss of trees, which slow the flow of run-off and promote infiltration, increases the risk of flooding.



Figure 6.1.4.*b*: Peak Downs and Saraji Coal Mine in Queensland, Australia. To access the coal, vegetation was removed, destroying habitat for native organisms. Image by Lock the Gate Alliance (CC-BY).



Figure 6.1.4.*c*: Mountaintop Removal Coal Mining in Martin County, Kentucky Photograph shows mountaintop coal removal mining in Martin County, Kentucky. Source: Flashdark.

Coal miners face health hazards such as explosions, mine collapse, and exposure to toxic fumes. **Black lung disease** is a respiratory condition characterized by coughing and shortness of breath that occurs in miners exposed to too much coal dust. Residents near mines also risk exposure to coal dust and underground toxins following explosions. As a result of exposure to toxins, birth defects and other health problems are common in residents near mines.

Coal is considered the “dirtiest” source of energy because its combustion results in the most air pollution. Coal power plants emits a variety of air pollutants including sulfur dioxide, nitrogen oxide, and heavy metals. Sulfur dioxide and nitrogen oxide are sources of acid rain (acid deposition), smog, and health issues. Heavy metals cause neurological and developmental problems in humans and other animals. Burning of coal emits particulate matter and higher amounts of carbon dioxide per unit of energy than the use of oil or natural gas. Carbon dioxide is the most frequently emitted greenhouse gas and causes climate change. In 2018, electricity generation was responsible 27% of greenhouse gas emissions in the United States, and much of this was emitted from coal power plants. The transportation of coal usually relies on fossil fuels, releasing further pollution.

Ash (including fly ash and bottom ash) is a residue created when coal is burned at power plants. In the past, fly ash was released into the air through the smokestack, where it would contribute to particulate matter air pollution. Laws now require that much of the fly ash now must be be captured by pollution control devices, like scrubbers. In the United States, fly ash is generally stored at

coal power plants or placed in landfills. Ash from storage or landfills can spill or leach into groundwater, resulting in water pollution.

Conventional Extraction Oil and Natural Gas

Exploring and drilling for oil degrades land and ocean habitats. On land, extensive infrastructure such as road networks, transport pipelines and housing for workers are needed to support a full-scale drilling operation. These can pollute soil and water, fragment habitats, and disturb wildlife. Extraction of oil and natural gas is also hazardous for workers, who have a high incidence of cancer and heart disease.

Human-caused oil spills in rivers and oceans harm ecosystems. From an economic perspective, oil spills disrupt the fishing industry and tourism. Oil spills at sea are generally much more damaging than those on land, since they can spread for hundreds of nautical miles in a thin oil slick which can cover beaches with a thin coating of oil. This can kill sea birds, mammals, shellfish and other organisms it coats. Oil spills on land are more readily containable if a makeshift earth dam can be rapidly bulldozed around the spill site before most of the oil escapes, and land animals can avoid the oil more easily.

Oil spills can result from supertanker accidents such as the Exxon Valdez in 1989, which spilled 10 million gallons of oil into the rich ecosystem of coastal Alaska and killed massive numbers of animals. The largest marine oil spill began in April 2010 when a natural gas explosion at an oil well 65 km offshore of Louisiana occurred on Deepwater Horizon Oil Rig. It killed 11 employees and flowed for 3 months in 2010, releasing an estimated 200 million gallons of oil (figure 6.1.4.*d*). Wildlife, ecosystems, and people's livelihood were adversely affected. A lot of money and huge amounts of energy were expended on immediate clean-up efforts. The long-term impacts are still not known. The National Commission on the Deepwater Horizon Oil Spill and Offshore Drilling was set up to study what went wrong. The worst oil spill ever occurred during the Persian Gulf war of 1991, when Iraq deliberately dumped approximately 200 million gallons of oil in offshore Kuwait and set more than 700 oil well fires that released enormous clouds of smoke and acid rain for over nine months.



Figure 6.1.4.*d*: Platform supply vessels battle the blazing remnants of the off shore oil rig Deepwater Horizon. A Coast Guard MH-65C dolphin rescue helicopter and crew document the fire aboard the mobile offshore drilling unit Deepwater Horizon, while searching for survivors. Multiple Coast Guard helicopters, planes and cutters responded to rescue the Deepwater Horizon's 126 person crew. Image and caption (modified) from US Coast Guard (public domain).

During an oil spill on water, oil floats to the surface because it is less dense than water, and the lightest hydrocarbons evaporate, decreasing the size of the spill but polluting the air. Then, bacteria begin to decompose the remaining oil, in a process that can take many years. After several months only about 15% of the original volume may remain, but it is in thick asphalt lumps, a form that is particularly harmful to birds, fish, and shellfish. Cleanup operations can include a variety of components, but each has its procs and cons. Skimmer ships that vacuum oil from the water surface, but these are effective only for small spills. Controlled burning works only in early stages before the light, ignitable part evaporates, but this also pollutes the air. **Dispersants** are detergents that break up oil to accelerate its decomposition, but some dispersants may be toxic to the ecosystem. **Bioremediation** refers to adding microorganisms that specialize in quickly decomposing oil, but this can disrupt the natural ecosystem.

Unconventional Extraction of Oil and Natural Gas

Fracking causes more environmental damage than conventional extraction. The considerable use of water (figure 6.1.4.*e*) may affect the availability of water for other uses in some regions, and this can affect aquatic habitats. In fact, fracking consumes more water than the use of nuclear energy, coal, or conventional oil and natural gas. If mismanaged, hydraulic fracturing fluid can be released by spills, leaks, or various other exposure pathways that contaminate land and groundwater (figure 6.1.4.*f*). Fracking fluid flowback – the fluid pumped out of the well and separated from oil and gas – not only contains the chemical additives used in the drilling process but also contains heavy metals, radioactive materials (which release radiation), volatile organic compounds, benzene (a carcinogen), toluene, ethylbenze, xylene, and other toxic air pollutants. Volatile organic compounds (VOCs) can react with the atmosphere to form ground-level ozone, which is associated with respiratory disease. Toulene can cause dizziness, confusion, headaches, and miscarriages. Ethylbenzene is a possible carcinogen that also causes dizziness, eye irritation, and hearing loss. Xylene also causes dizziness and headaches and furthermore can be fatal at high concentrations. In some cases, this contaminated water is sent to water treatment plants that are not equipped to deal with some of these classes of contamination. Finally, injecting wastewater for disposal can even induce earthquakes.

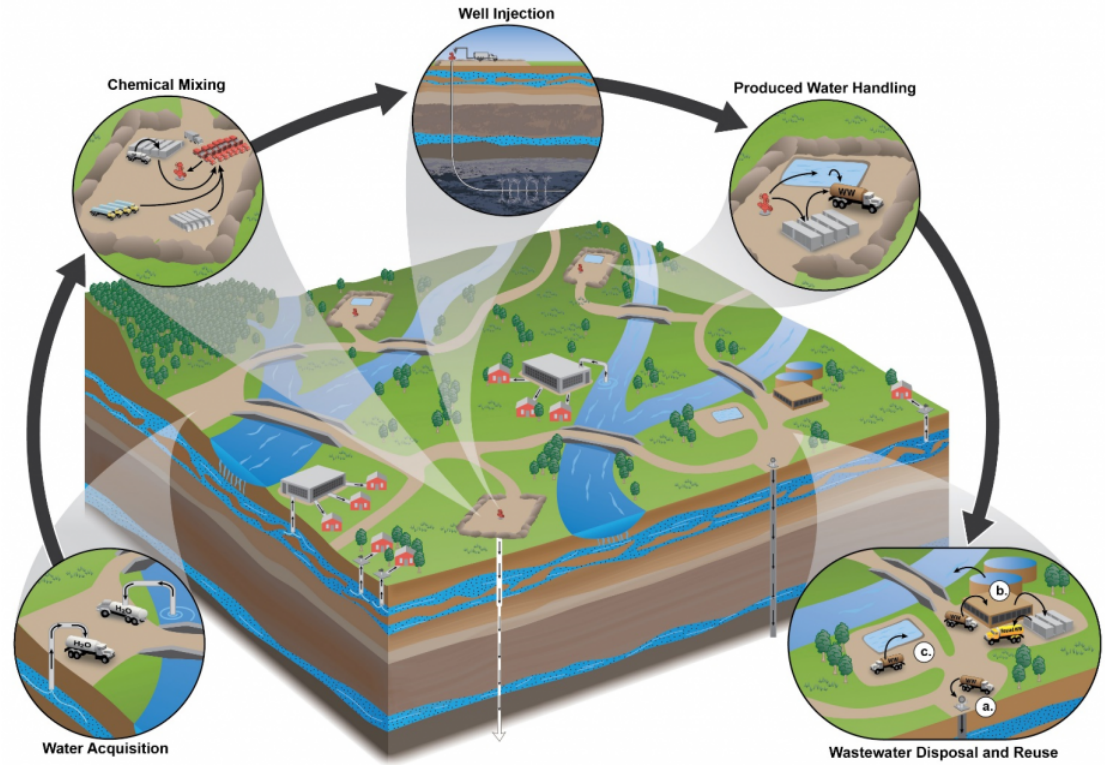


Figure 6.1.4.*e*: Fracking influences the water cycle. Water acquisition is the withdrawal of groundwater or surface water to make hydraulic fracturing fluids. Chemical mixing involves combining a base fluid, sand, and additives at the well site to create hydraulic fracturing fluids. During well injection, hydraulic fracturing fluids move through the oil and gas production well and in the targeted rock formation. Produced water handling refers to the onsite collection and handling of water that returns to the surface after hydraulic fracturing and the transportation of that water for disposal or reuse. Finally, the disposal and reuse of wastewater occurs. If this is not done properly, wastewater can pollute surrounding areas. Image and caption (modified) by EPA (public domain).

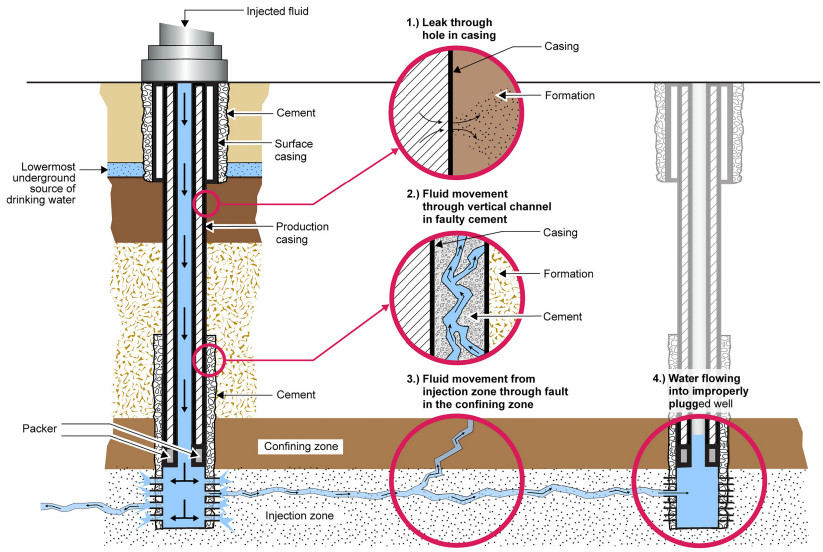


Figure 6.1.4.*f*: The toxic fluid injected during fracking can escape, contaminating groundwater. For example, it can leak through a hole in the casing, move through faulty cement, or move through a fault in the rock that confines the deposit. Image by Government Accountability Office (public domain).

Other unconventional sources of fossil fuels can also harm the environment. Surface mining of tar sands or oil shales requires the removal of all vegetation and leaves pollutants behind, causing habitat loss (figure 6.1.4.*g*).



Figure 6.1.4.*g*: An open pit mine in the tar sands oil fields of Alberta, Canada. Conifer trees, which previously provided habitat to native animals, have been cleared to allow for mining. Image by Howl Arts Collective (CC-BY).

Transportation, Refineries, and Combustion

Natural gas is released into the atmosphere from coal mines, oil and gas wells, and natural gas storage tanks, pipelines, and processing plants. These leaks are the source of about 25% of total U.S. methane emissions, which translates to three percent of total U.S. greenhouse gas emissions, contributing to climate change. When natural gas is produced but cannot be captured and transported economically, it is "flared," or burned at well sites, which converts it to carbon dioxide. This is considered to be safer and better than releasing methane into the atmosphere because carbon dioxide is a less potent greenhouse gas than methane. However, when natural gas with high concentrations of the toxic gas hydrogen sulfide is flared, it produces carbon dioxide, carbon monoxide, sulfur dioxide, nitrogen oxides, and many other compounds (see Air Pollution for more details).

Leaks also happen when we use petrochemicals on land. For example, gasoline sometimes drips onto the ground when people are filling their gas tanks, when motor oil gets thrown away after an oil change, or when fuel escapes from a leaky storage tank. When it rains, the spilled petrochemicals get washed into the gutter and eventually flow to rivers and into the ocean. Another way that oil sometimes gets into water is when fuel is leaked from motorboats and jet skis. When a leak in a storage tank or pipeline occurs,

petrochemicals can also get into the ground, and the ground must be cleaned up. To prevent leaks from underground storage tanks, all buried tanks are supposed to be replaced by tanks with a double lining.

Oil refining emits a variety of toxins and is the single largest source of benzene (figure 6.1.4.*g*). As a result, residents living near oil refineries have a high incidence of cancer, asthma, and birth defects. When petrochemicals such as gasoline or diesel are burned, they release a variety of air pollutants, including carbon dioxide (a cause of climate change), sulfur dioxide, nitrous oxides, volatile organic compounds (VOCs), particulate matter, and lead (see Air Pollution for more details). The transport of oil by ship or trunk also requires energy in the form of fossil fuels, generating more pollutants. Compared to oil and coal, burning natural gas releases the fewest pollutants and greenhouse gases.



Figure 6.1.4.*g*: Pollutants are emitted from the Pine Bend oil refinery in Rosemount, Minnesota. Image by Tony Webster (CC-BY).

Solutions

Reclamation can mitigate the habitat damage that results from mining or extracting fossil fuels. It involves restoring the land to an extent after mining or extraction is complete. This can entail returning displaced land and covering with top soil, which protects organisms from heavy metals, radioactive materials, and other underground toxins. Additionally, acids, which often form from the leaching of sulfates from underground rocks, may be neutralized. Vegetation is then planted, and water flow if disrupted is somewhat restored. Of course, the intricate topography, network of streams, and mature vegetation (such as large trees in forest) that may have been present prior to mining cannot be recreated, but reclamation makes it easier for native species to begin recolonizing the area.

Clean coal technologies can limit the air pollution released when burning coal. Some of these technologies remove toxins from coal before burning it while others capture toxins that are released while burning coal. For example, **smokestack scrubbers** in power plants clean sulfur dioxide, nitrous oxide, particulate matter, and mercury from the smoke before it is released. **Carbon capture and sequestration** involves capturing carbon dioxide released and storing it, but it requires 25-40% more energy, reducing the efficiency of coal (figure 6.1.4.*h*). In this process, smoke from a coal power plant is passed through a solvent to trap carbon dioxide, but other waste gases are still released in the smoke. Carbon dioxide is then separated from the solvent. Some can be used in industry (such as for carbonated beverages or to tertiary recovery of oil), and the rest is sequestered (stored) underground. Note that clean coal technologies can reduce coal's contribution to climate change and reduce the amount of toxins that are released, but it does not fully prevent coal-generated air pollution (figure 6.1.4.*i*).

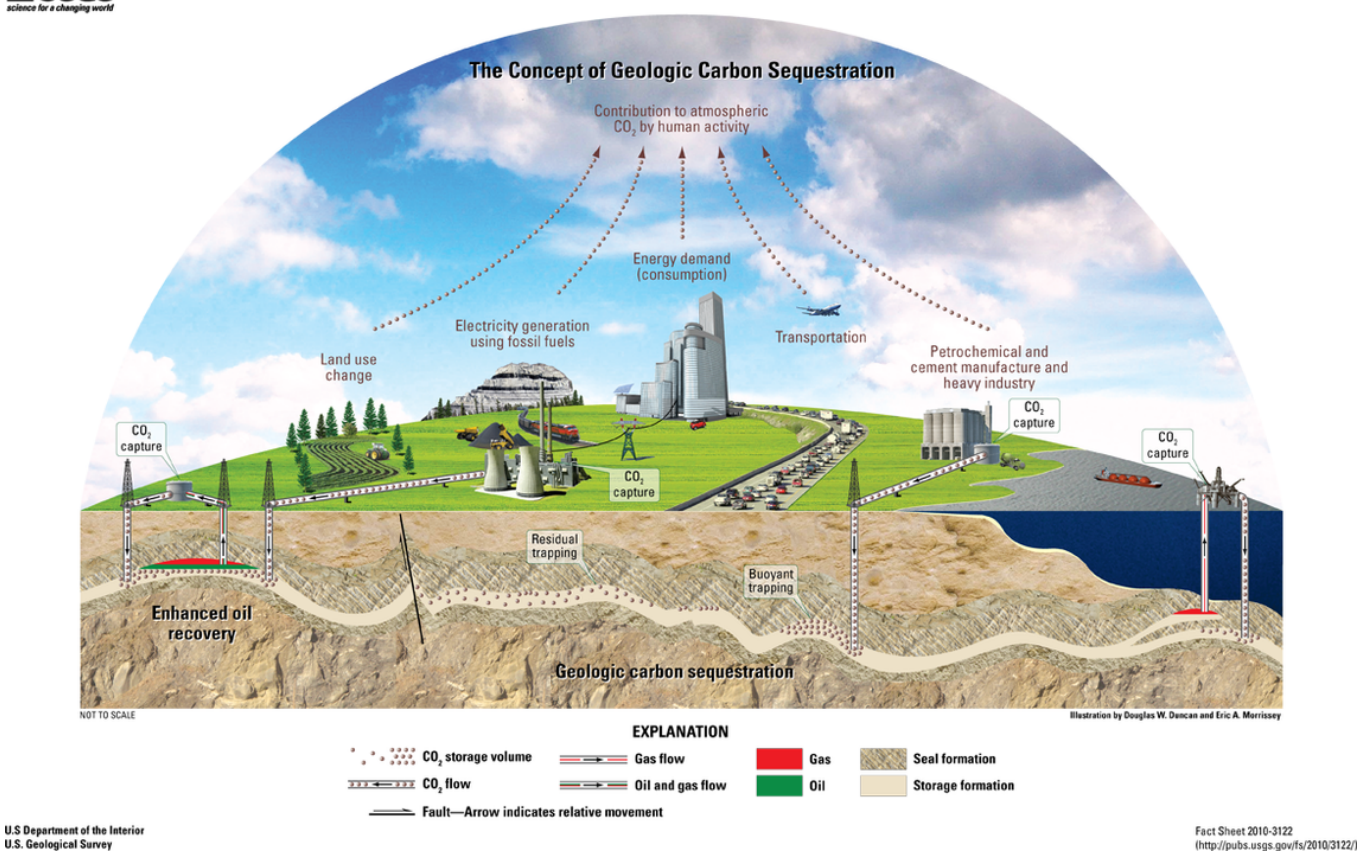


Figure 6.1.4.*h*: Burning fossil fuels releases carbon dioxide into the atmosphere, but there are also opportunities to sequester (store) carbon dioxide, offsetting some fo this damage. Some of the carbon dioxide injected during tertiary recovery (enhanced oil recovery) remains underground. Additionally, carbon dioxide may be captured from smoke released by power plants and other industries and stored under ground. Image by USGS (public domain).



Figure 6.1.4.*i*: A protest against clean coal at the United Nations Climate Change Conference (COP19) in 2013. Image by 350.org (CC-BY-NC-SA).

Because fossil fuels are nonrenewable, reserves will eventually be depleted, and the world will need to fully rely on other energy sources. Those concerned about the environmental and health consequences of fossil fuels advocate for making this transition as

soon as possible. This is because the technologies and practices discussed above do not fully prevent fossil fuels from causing environmental damage and causing health hazards for workers and the general public. The next two chapters discuss nuclear energy and renewable energy, which are alternatives to fossil fuels. As even these alternatives have their disadvantages, **energy conservation** (using energy more efficiently and limiting unnecessary energy use) is also critical.

Attribution

Modified by Melissa Ha from the following sources:

- Challenges and Impacts of Energy Use, Non-Renewable Energy Sources, and Water Pollution from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Fossil Fuels from *An Introduction to Geology* by Johnson et al. (licensed under CC-BY-NC-SA)
- Chapter 4: Non-Renewable Energy from *Introduction to Environmental Science: 2nd Edition* (2018) Biological Sciences Open Textbooks by Zehnder, Caralyn; Manoylov, Kalina; Mutiti, Samuel; Mutiti, Christine; VandeVoort, Allison; and Bennett, Donna (licensed under CC-BY-NC-SA).

This page titled 6.1.4: Consequences of Fossil Fuels is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **16.4: Consequences of Fossil Fuels** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

6.1.5: Data Dive- Global Fossil Fuel Consumption

Overview

Our World in Data (OWID) is a scientific online publication that focuses on using research and data to help tackle global concerns such as poverty, disease, hunger, climate change, war, and inequitable treatment of our world's most vulnerable and unstable communities. Their website is particularly known for publishing a variety of graphs, some even interactive, to present the research that helps explain causes and consequences of global concerns to the public. One example graph, seen below, illustrates global fossil fuel consumption:

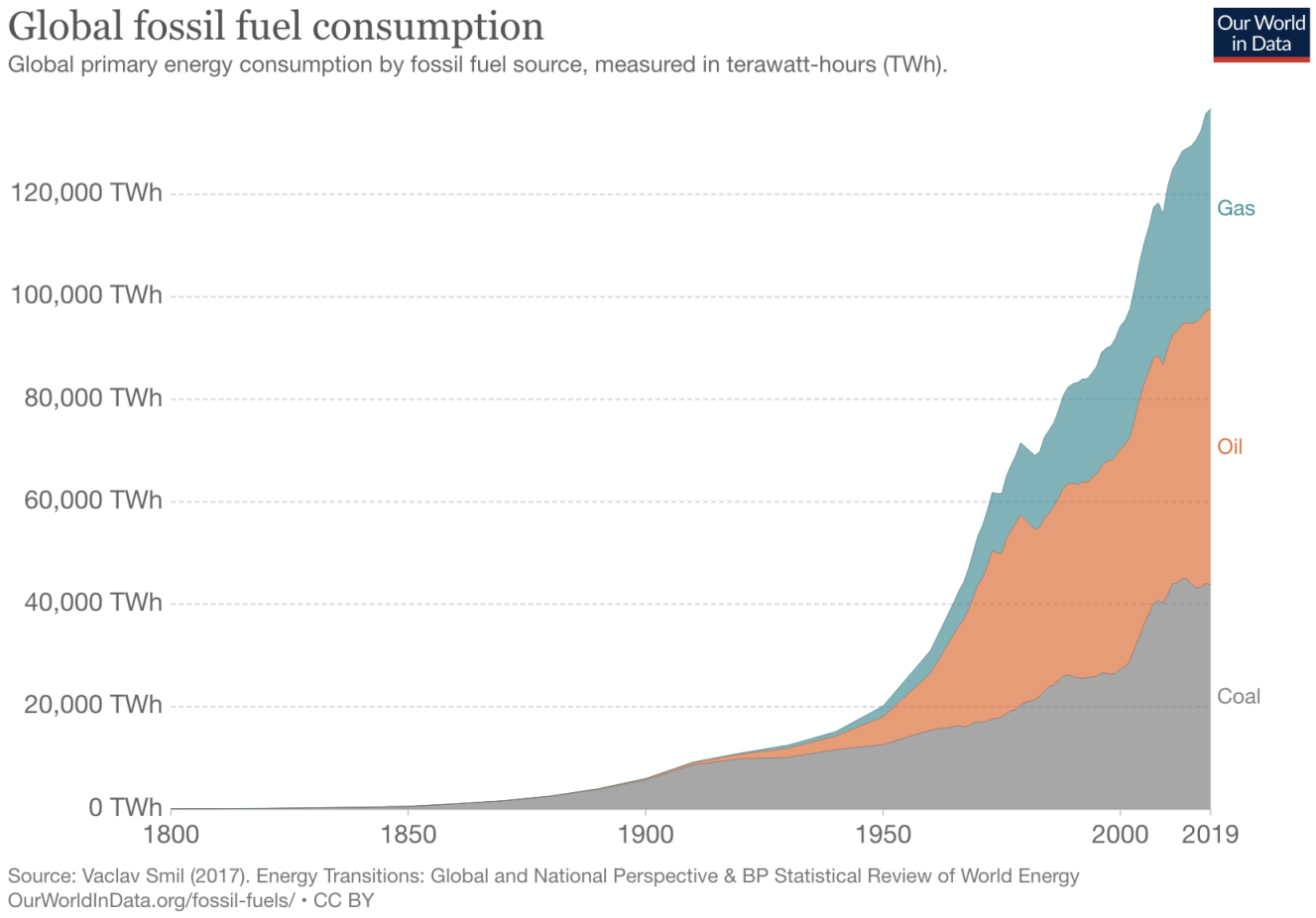


Figure 6.1.5.*a*: Global consumption of gas, oil, and coal. Graph by Our World in Data (CC-BY)

Questions

1. What kind of graph is this?
2. What question(s) are the authors trying to answer with this graph?
3. What are the observed patterns of fossil fuel use (gas, oil, and coal) in this graph?
4. Make a prediction for each fossil fuel use pattern for the next 50 years. For each prediction, state why you think that pattern will occur.
5. How do you think climate scientists are using a graph like this to predict climate into the future?

Raw Data For Above Graph(s)

Raw data table link to Our World in Data graph above. (CC-BY)

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 6.1.5: Data Dive- Global Fossil Fuel Consumption is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **16.5: Data Dive- Global Fossil Fuel Consumption** by Melissa Ha and Rachel Schleiger has no license indicated.

6.1.6: Review

Summary

After completing this chapter you should be able to...

- Distinguish among coal, oil, and natural gas and explain how each is formed.
- Describe the processes of mining, processing, transporting, and burning coal to produce electricity.
- Compare the processes of conventional oil and natural gas extraction and fracking.
- Identify the relative contributions of coal, oil, and natural gas to electricity generation in the U.S. and total energy consumption globally.
- Detail the advantages and disadvantages of fossil fuel use.
- Describe strategies to mitigate harm caused by fossil fuel use.

Fossil fuels are nonrenewable energy sources that were formed from ancient organisms. **Coal** is a solid fossil fuel formed from ancient swamp vegetation. **Oil** (petroleum) and natural gas both formed from marine microorganisms and are often found together; however, oil is a liquid fossil fuel, and natural gas is in gaseous form.

Coal is produced through **surface** and **subsurface mining**. Some oil and natural gas reserves can be recovered through **conventional extraction**, but unconventional reserves are recovered through **hydraulic fracturing** (tight oil and natural gas; tight sands) or mining (oil shale and tar sands). To produce electricity from coal or natural gas, the fuel is combusted (burned). The heat is used to produce steam, and the pressure from the steam turns a turbine, which powers a generator. **Crude oil** must be refined to separate it into a variety of **petrochemicals**, each with a different use.

Nationally and globally, fossil fuel consumption remains high. **Proven reserves** of oil and natural gas are projected to last another 50 years while coal reserves are projected to last another 115 years. The U.S. has a large supply of the world's coal but relies largely on imports to meet its oil demands. The world's oil supply is concentrated in OPEC countries, which have great influence over global rates of production.

While existing technologies facilitate continued fossil fuel use, which support local and national economies, fossil fuels cause a variety of health and environmental consequences. Workers and those living near mines or oil refineries are at the greatest risk of health consequences, but combustion of fossil fuels releases air pollutants that harm the general public and native organisms. Furthermore, mining, drilling, and fracking, destroy habitats and release pollutants. **Clean coal technologies** and **reclamation** reduces but does not eliminate these negative consequences.

Attribution

Melissa Ha (CC-BY-NC)

This page titled 6.1.6: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **16.6: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

6.2: Nuclear Energy

Chapter Hook

It's like a scene from a post-apocalyptic fictional novel. A city that is slowly deteriorating year after year with human absence. Schools, homes, and business with doors wide open, belongings strewn all over floors and tables, indicating a serious panic and rush from people that once called this city theirs. Unfortunately, Chernobyl is not a fictional novel. It was a city in the former USSR that is the site of the worst nuclear accident in history. On April 25th and 26th in 1986 one of the nuclear reactors exploded releasing up to 30% of Chernobyl's 190 metric tons of uranium into the atmosphere. Currently, there is a 19-mile-wide prohibition zone around the epicenter of the disaster, and it is estimated not to be safe for 20,000 years! This is the terrible price to pay for the sake of electricity.



Figure 6.2.*a*: Chernobyl radiation map. Image Sting (vectorisation), MTruch (English translation), Makeemlighter (English translation) in Wikimedia Commons (CC-BY-SA2.5)

Nuclear power is energy released from the radioactive decay of elements, such as uranium, which releases large amounts of energy. It generally refers to using the heat energy released from **nuclear fission reactions** to produce electricity. Nuclear power plants produce no carbon dioxide and, therefore, are often considered an alternative fuel (fuels other than fossil fuels)**.**

6.2.1: Radioactive Isotopes

6.2.2: Generating Electricity with Nuclear Energy

6.2.3: Nuclear Energy Consumption

6.2.4: Consequences of Nuclear Energy

6.2.5: Data Dive- Global Nuclear Power Generation

6.2.6: Review

Attribution

Modified by Melissa Ha and Rachel Schleiger from Non-Renewable Energy Sources from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 6.2: Nuclear Energy is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

6.2.1: Radioactive Isotopes

Recall that an **atom** is the smallest component of an element that retains all of the chemical properties of that element (see Matter). As previously discussed, atoms contain uncharged **neutrons** and positively charged **protons** in the nucleus. Negatively charged **electrons** surround the nucleus. The **atomic mass** of an atom is determined by the number of protons and neutrons because the mass of electrons is negligible. Each proton or neutron weighs 1 atomic mass unit (AMU). The atomic mass values displayed in the periodic table of elements are not whole numbers because they represent the average atomic mass for atoms of that element (figure 6.2.1.*a*). Atoms of the same element do not necessarily have the same mass because they can differ in neutron number.

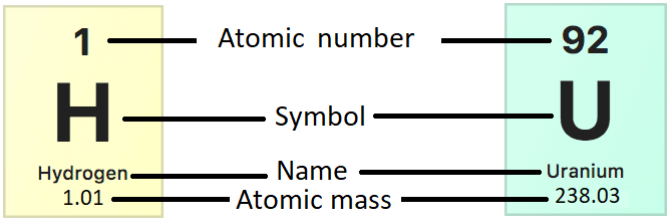


Figure 6.2.1.*a*: The whole number in each cell of periodic table is the atomic number, or number of protons (1 for hydrogen and 92 for uranium). The atomic mass number is also shown, which is the average number of protons and neutrons (1.01 for hydrogen and 238.03 for uranium). Images cropped and labeled from National Center for Biotechnology Information (public domain).

Isotopes are different forms of the same element that have the same number of protons, but a different number of neutrons. Some elements, such as carbon, potassium, and uranium, have naturally occurring isotopes. Carbon-12, the most common isotope of carbon, contains six protons and six neutrons. Therefore, it has a mass number of 12 (six protons and six neutrons) and an atomic number of 6 (which makes it carbon). Carbon-14 contains six protons and eight neutrons. Therefore, it has a mass number of 14 (six protons and eight neutrons) and an atomic number of 6, meaning it is still the element carbon. These two alternate forms of carbon are isotopes. Some isotopes are unstable and emit radiation in the form of particles and energy to form more stable elements. Some forms of radiation are dangerous. These are called **radioactive isotopes** or radioisotopes (figure 6.2.1.*b*). During **radioactive decay**, one type of atom can change into another type of atom in this way (figure 6.2.1.*c*).

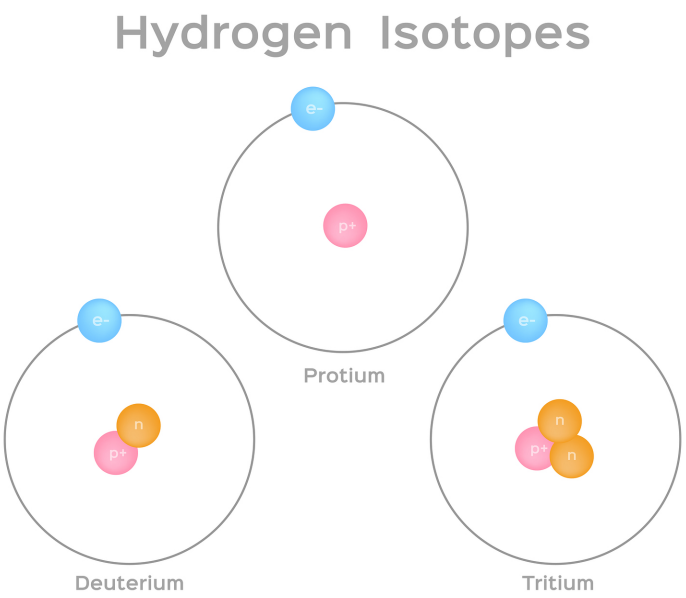


Figure 6.2.1.*b*: Isotopes of hydrogen. All of these atoms have one proton (pink circle labeled "p^+"), but protium has no neutrons, deuterium has one neutron (orange circle labeled "n"), and tritium has two neutrons. The proton and neutron(s) are located in the center of the atom (the nucleus). An electron (blue circle labeled "e^-") rotates around each atom nucleus. Image from National Isotope Development Center/U.S. Department of Energy Isotope Program (public domain).

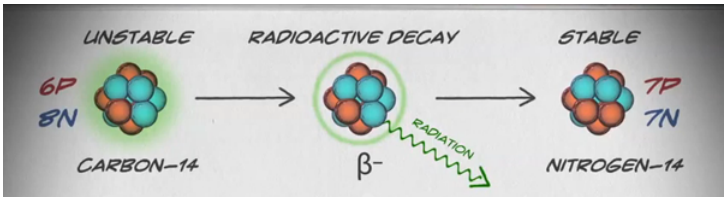


Figure 6.2.1.*c*: A **radioactive isotope** of carbon (carbon-14) has six protons and eight neutrons. It decays to a stable isotope of nitrogen (nitrogen-14), which has seven protons and seven neutrons. Radioactive decay releases radiation. (The particular type of radiation that occurs in this example is called beta minus decay, β-.) Carbon-14 decays at a predictable rate, with half of it decaying every 5730 years. Because carbon is abundant in organisms, this predictable decay rate is commonly used for dating fossils. Image from CDC (public domain).

Half-Life

The **half-life** is the amount of time that it takes for half of the original radioactive isotope to decay (figure 6.2.1.*d*). For example, the half-life of uranium-238 is about 4.5 billion years. After 4.5 billion years, only half (50%) of the original amount of uranium-238 will remain. The rest will have decayed to thorium-234 (which is also radioactive and quickly decays to a series of radioactive isotopes, until it ultimately becomes lead-206, which is stable; figure 6.2.1.*e* – *f*). After two half-lives (9 billion years), only half of the 50% would remain (25% of the original). After three half-lives, only 12.5% of the original uranium-238 would remain.

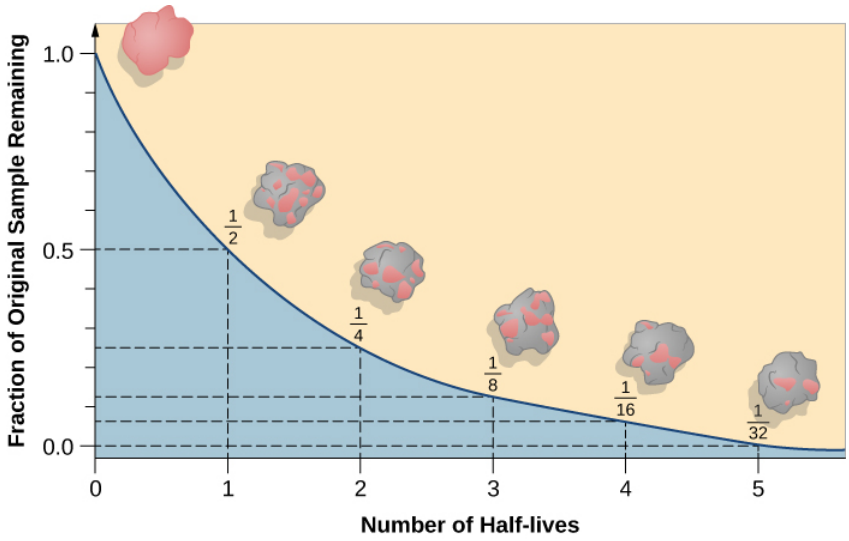


Figure 6.2.1.*d*: After each half life, 50% of a radioactive isotope decays. After one half-life, 50% (1/2; 0.5) of a radioactive isotope remains. After two half-lives, only 25% of the original radioactive isotope remains. After three half-lives, it is 12.5%; four half-lives = 6.25%; five half-lives = 3.125%. Image by Fraknoi, Morrison, and Wolff/OpenStax (CC-BY). Access for free at openstax.org.

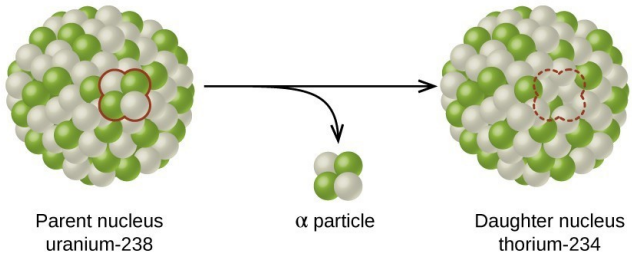


Figure 6.2.1.*e*: When unstable uranium-238 decays, it emits an alpha (α) particle, which is two protons and two neutrons. This changes it into a new element (thorium-234). Image by OpenStax (CC-BY). Download for free at openstaxcollege.org.

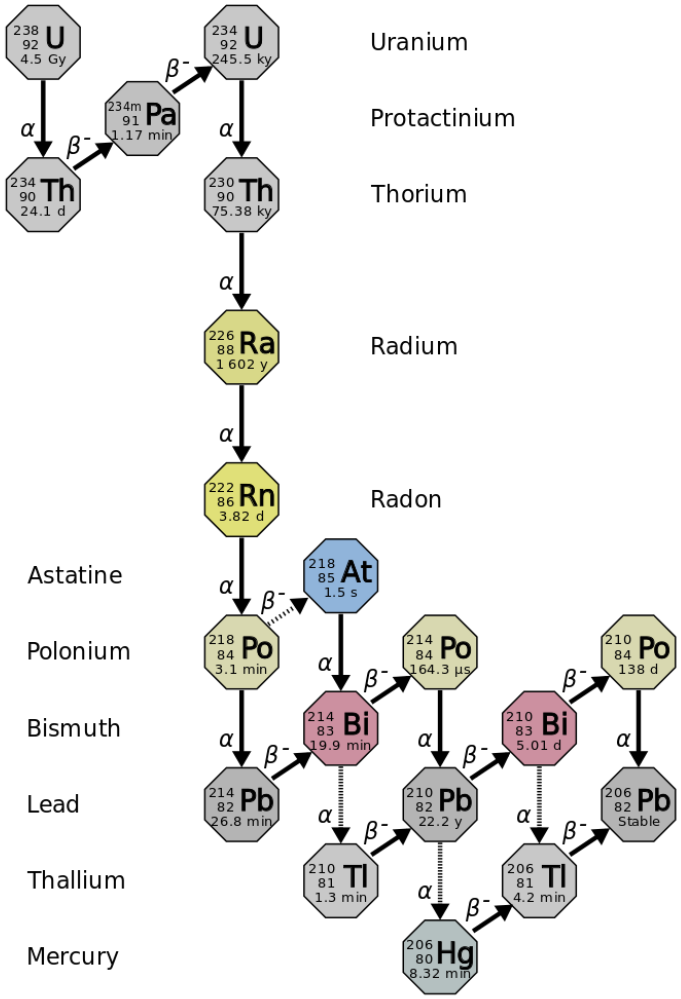


Figure 6.2.1.f: The decay chain of uranium-238. Each isotopes mass, atomic number, and half-live is listed in order, to the right of the atomic symbol. Each arrow is labeled with the type of radiation released: alpha (α) radiation, which is two protons and two neutrons, or beta (β) radiation, which is a high-energy electron. Decay continues from uranium until ending at a stable (non-radioactive) isotope, lead (Pb-206). Image by Tosaka (CC-BY).

Evolution in Action: Carbon Dating

Carbon-14 (^{14}C) is a naturally occurring radioisotope that is created in the atmosphere by cosmic rays. This is a continuous process, so more ^{14}C is always being created. As a living organism develops, the relative level of ^{14}C in its body is equal to the concentration of 14C in the atmosphere. When an organism dies, it is no longer ingesting ^{14}C, so the ratio will decline. ^{14}C decays to ^{14}N by a process called beta decay; it gives off energy in this slow process (figure 6.2.1.c). After approximately 5,730 years, only one-half of the starting concentration of ^{14}C will have been converted to ^{14}N. The time it takes for half of the original concentration of an isotope to decay to its more stable form is called its half-life.

Because the half-life of ^{14}C is long, it is used to age formerly living objects, such as fossils. Using the ratio of the ^{14}C concentration found in an object to the amount of ^{14}C detected in the atmosphere, the amount of the isotope that has not yet decayed can be determined. Based on this amount, the age of the fossil can be calculated to about 50,000 years (figure 6.2.1.g below). Isotopes with longer half-lives, such as potassium-40, are used to calculate the ages of older fossils. Through the use of carbon dating, scientists can reconstruct the ecology and biogeography of organisms living within the past 50,000 years.



Figure 6.2.1.*g*: The age of remains that contain carbon and are less than about 50,000 years old, such as this pygmy mammoth, can be determined using carbon dating. (credit: Bill Faulkner/ NPS)

Nuclear Fission Reactions

Nuclear fission reactions are those that involve splitting the nucleus of an atom (figure 6.2.1.*h*). They can be induced by blasting radioactive elements with neutrons. As with natural radioactive decay, induced nuclear fission reactions release energy. The heat energy released when nuclear fission can be used to generate electricity. This is the basis of nuclear power. Currently, **uranium-235** (^{235}U; an isotope of uranium with an atomic mass of 235) is currently used as fuel for nuclear fission reactions (figure 6.2.1.*h*).

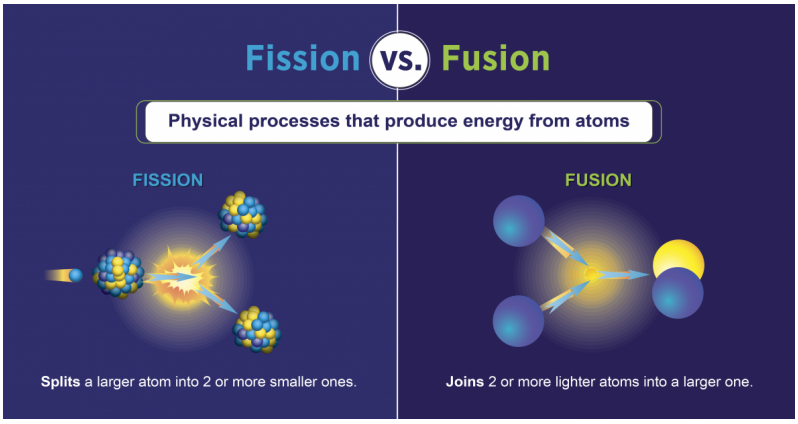


Figure 6.2.1.*h*: Nuclear fission and fusion are physical processes that produce energy from atoms. Nuclear fission involves splitting the nucleus of an atom, but nuclear fusion involves combining smaller nuclei into a larger one. Image by Sarah Harmann/U.S. Department of Energy (public domain).

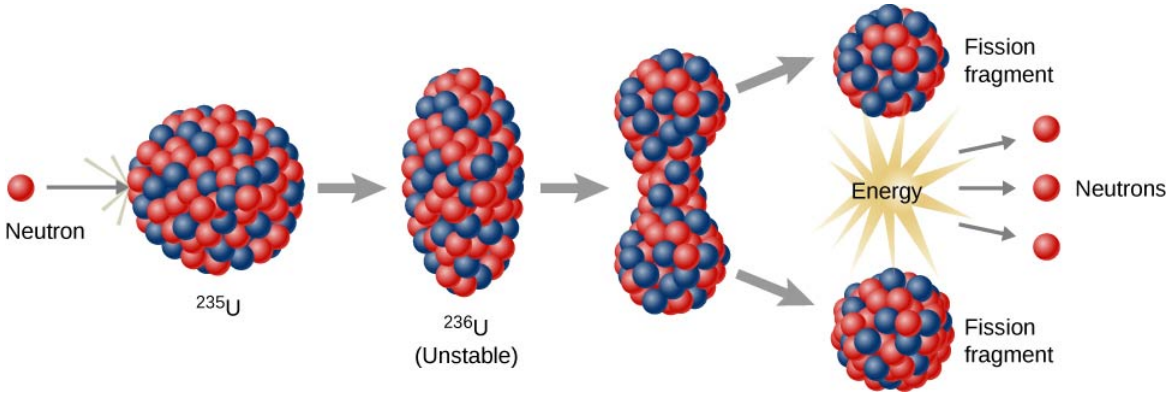


Figure 6.2.1.*i*: Fission of uranium-235 (^{235}U) can be induced by bombarding it with a neutron. This produces U-236, which is unstable and splits into fission fragments and additional neutrons. Energy is released as a result. Red spheres are neutrons, and blue spheres are protons. Image by BC Open Textbooks (CC-BY).

Attribution

Modified by Melissa Ha from Matter from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 6.2.1: Radioactive Isotopes is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **17.1: Radioactive Isotopes** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

6.2.2: Generating Electricity with Nuclear Energy

The Nuclear Fuel Cycle

Uranium ore must be mined, milled, and enriched to produce nuclear fuel. The **nuclear fuel cycle** represents the progression of nuclear fuel from creation to disposal (figure 6.2.2.*a*). The first stage of the nuclear fuel cycle is uranium recovery, in which uranium ore is **mined**. It is then milled to produce yellowcake (uranium ore concentrate/uranium oxide/U_3O_8). **Milling** separates the uranium from the other parts of the ore. Each ton of mined uranium ore typically yields 1-4 pounds of yellowcake (0.05% to 0.20% yellowcake). Next, the uranium ore concentrate is converted into uranium hexafluoride (UF_6). It is then **enriched** to increase the concentration of uranium-235 (^{235}U) relative to ^{238}U. During **fuel fabrication,** natural and enriched UF_6 is converted into into uranium dioxide (UO_2) or uranium metal alloys for use as fuel for nuclear power plants. Disposal of spent fuel rods and other hazardous waste generated in this process are discussed in Consequences of Nuclear Energy.

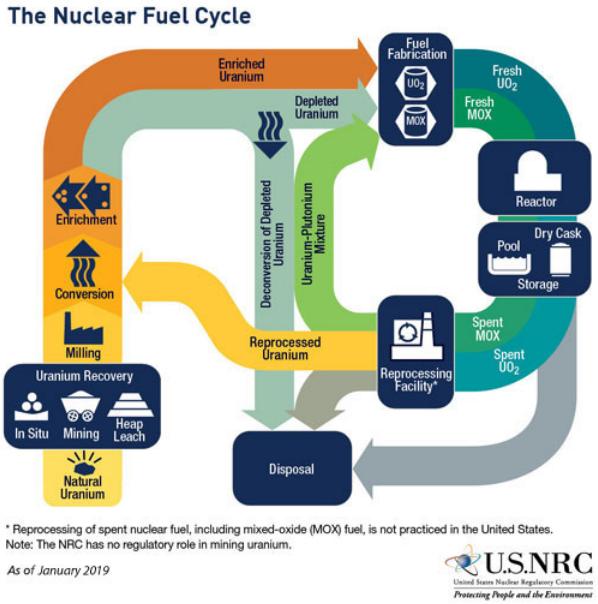


Figure 6.2.2.*a*: The nuclear fuel cycle begins with recover (mining) of natural uranium, followed by milling, conversion, and enrichment. The enriched uranium then undergoes fuel fabrication, producing uranium dioxide (UO_2) or uranium metal alloys (MOX). The fuel then goes to a nuclear reactor at the power plant. Spent fuel rods are stored in a pool or dry cask, potentially reprocessed, and sent to disposal. Reprocessed uranium can reenter earlier steps of the fuel cycle. Wastes produced during enrichment undergo deconversion before disposal. Image by U.S. NRC (public domain)

Nuclear Reactors

The fuel, which is now in the form of cylindrical ceramic pellets are then sealed into long metal tubes called **fuel rods**, which are assembled in reactor cores along with control rods. Each fuel pellet, which is about 1 cm in length, stores the same amount of energy as a ton of coal. Thousands of fuel rods form the **reactor core**, the site of nuclear fission in a nuclear power plant (figure 6.2.2.*b*).

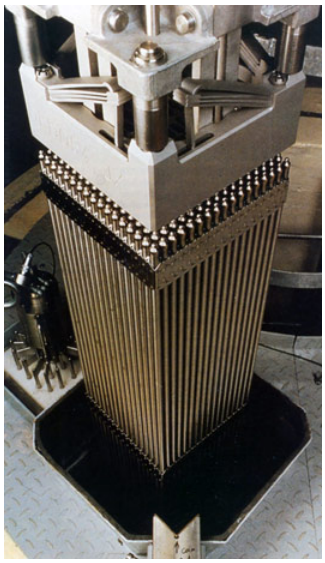


Figure 6.2.2.*b*: An assembly of fuel rods containing pellets of nuclear fuel. Image by Alternative Energies and Atomic Energy Commission, France (public domain)

Heat is produced in a nuclear reactor when neutrons strike uranium atoms, causing them to split in a continuous chain reaction that releases heat energy (figure 6.2.2.*c*). Specifically, fission of ^{235}U, releases additional neutrons, which then cause the fission of nearby ^{235}U nuclei. However, if fission occurs in too many atoms simultaneously, too much energy is released, which can cause an explosion or meltdown. This is prevented by **control rods**, which are made of a material such as boron that absorbs excess neutrons released in nuclear fission. When the neutron-absorbing control rods are pulled out of the core, more neutrons become available for fission, and the chain reaction speeds up, producing more heat. When they are inserted into the core, fewer neutrons are available for fission, and the chain reaction slows or stops, reducing the heat generated.

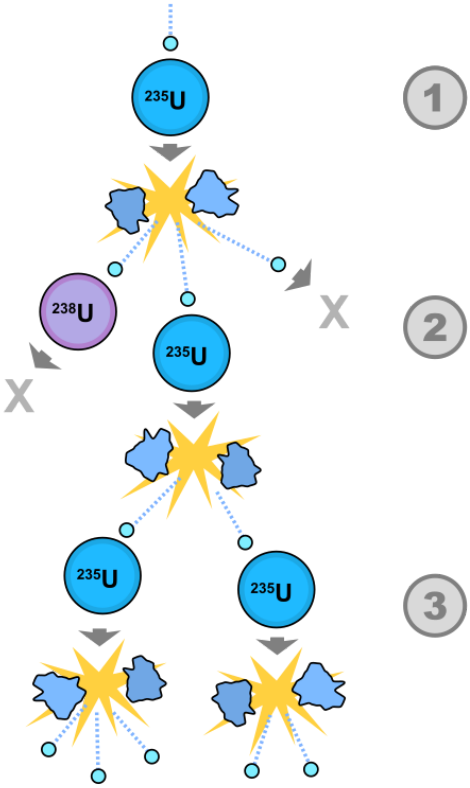


Figure 6.2.2.*c*: Schematic diagram of a fission chain reaction. (1) A uranium-235 atom absorbs a neutron, and fissions into two new atoms (fission fragments), releasing three new neutrons and energy. (2) One of those neutrons is absorbed by an atom of uranium-238, and does not continue the reaction. Another neutron is simply lost and does not collide with anything, also not continuing the reaction. However one neutron does collide with an atom of uranium-235, which then fissions and releases two neutrons and some binding energy. (3) Both of those neutrons collide with uranium-235 atoms, each of which fission and release between one and three neutrons, and so on. Image and caption (modified) by Fastfission (public domain).

Nuclear reactors (figure 6.2.2.*d*) contain the reactor core and the machinery needed to generate electricity from the heat released. The reactor core is submerged in water. In addition to transferring heat energy, the water also serves to slow down, or "moderate" the neutrons which is necessary for sustaining the fission reactions. Ultimately, the heat energy is used to generate high-pressure steam, which turns a turbine to generate electricity. The mechanism is similar to that of coal- or natural gas-generated electricity, but nuclear fission rather than combustion of coal is the source of heat energy.

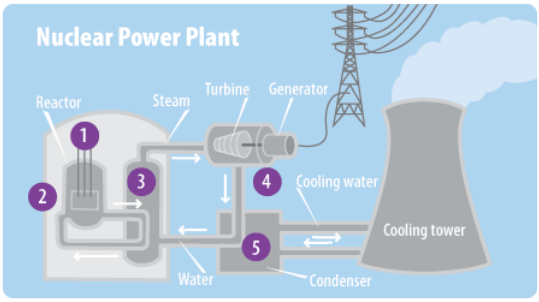


Figure 6.2.2.*d*: (1) In a nuclear reactor, fuel rods full of uranium pellets are placed in water. (2) Inside the fuel rods, uranium atoms split, releasing energy. (3) This energy heats water, creating steam. (4) The steam moves through a turbine, which turns a generator to create electricity. (5) In the condensor, the steam cools back into water, which can then be used over again. At some nuclear power plants, extra heat is released from a cooling tower. Image and caption (modified) by EPA (public domain).

There are two main types of nuclear reactors: pressurized water reactors and boiling water reactors.

Pressurized Water Reactor

In a **pressurized water reactor**, there are three separate streams of water: the water in contact with the reactor core, the water that produces steam, and the cooling water (figure 6.2.2.e). The reactor core is submerged in water, which is held by a steel vessel. This is surrounded by a containment structure. As the nuclear fission reaction heats the water surrounding it, the water is pumped in a cyclical stream. It transfers heat to the second stream of water, which is in a separate vessel. This second stream is kept at a lower pressure, allowing the water to boil and create steam The steam turns a turbine, generating electricity. The steam then is cooled in the **condenser** by a separate stream of cooling water. Because water from the reactor core does not mix with other parts of the reactor, not all of the reactor is radioactive.

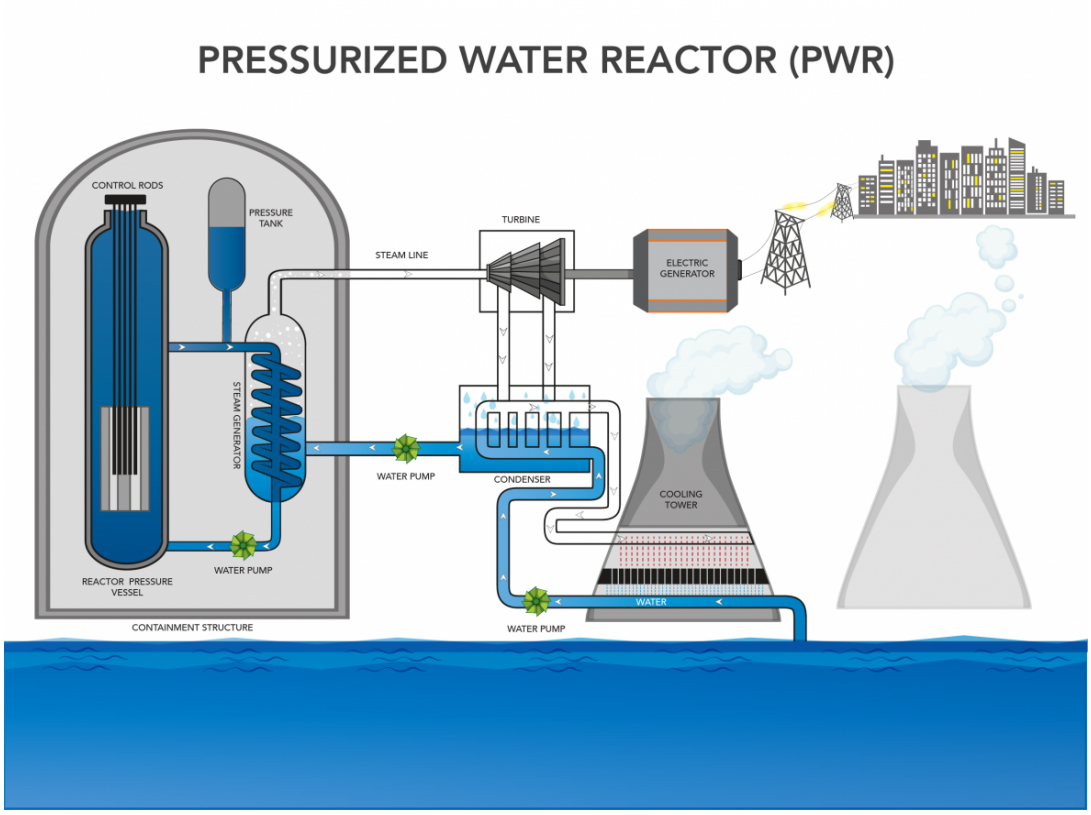


Figure 6.2.2.e: There are three separate streams of water in a pressurized water reactor. The first stream is associated with the reactor core in the reactor pressure vessel. The second stream is associated with the steam generator. The third stream is from an external source and is used to condense the steam. The reactor pressure vessel contains control rods and fuel rods (not labeled). The pressure tank regulates water pressure. Hot water from the reactor moves to the steam generator, where it heats water in the second stream to produce steam. Steam from the steam generator moves through the steam line and turns a turbine, powering the electric generator. Steam cools in the condenser, and the cooled water is then pumped back into the steam generator. The reactor pressure vessel, pressure tank, and steam generator are all in a containment structure. Water from an external source is used to cool steam in the condenser. The water from the external source cools in the cone-shaped cooling tower. Image by Office of Nuclear Energy/ U.S. Department of Energy (public domain).

Boiling Water Reactor

In a **boiling water reactor**, there are two separate streams of water: the water in contact with the reactor core and the cooling water (figure 6.2.2.f). The reactor core heats the water in which it is submerged. This water is held by a steel vessel that is surrounded by a containment structure. The steam produced as the reactor core heats water turns a turbine, which generates electricity. The steam then is cooled in the condenser by a separate stream of cooling water. Because water from the reactor core comes in contact with all parts of the reactor, the entire thing is radioactive.

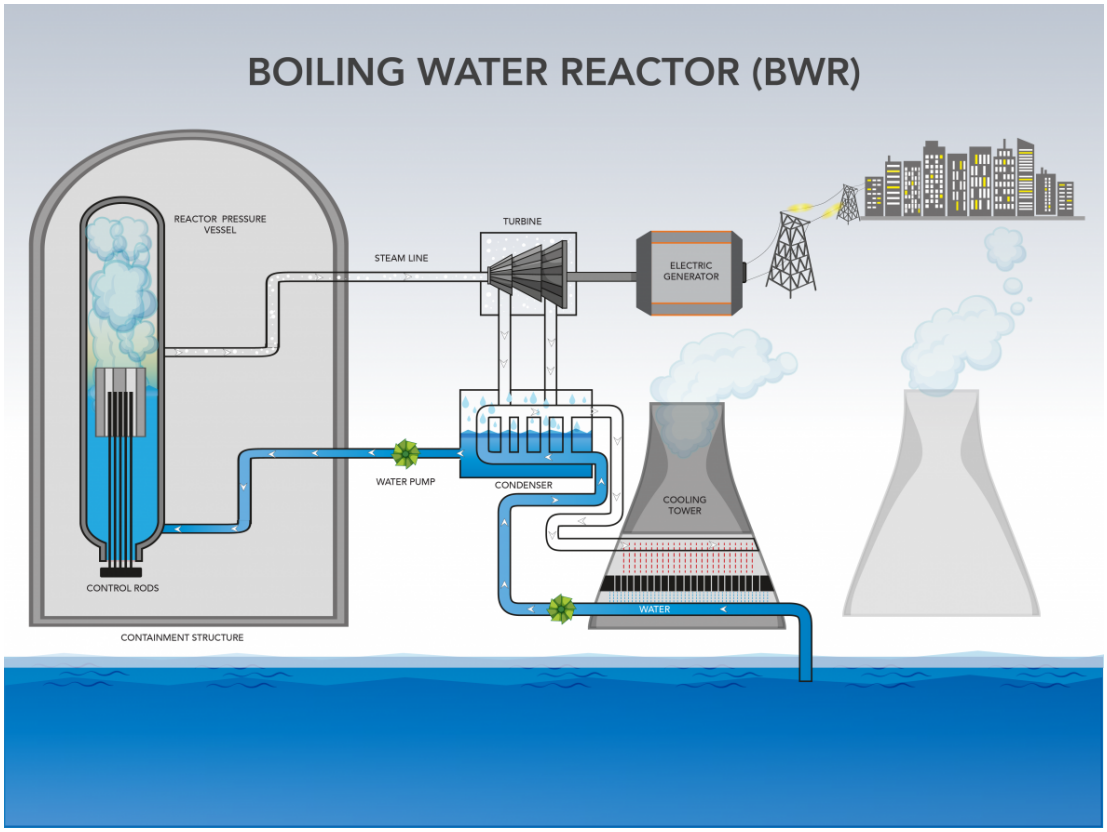


Figure 6.2.2.f: A boiling water reactor has two streams of water. In the first stream, the same water that bathes the reactor core in the reactor pressure vessel and turns to steam and turns a turbine. The third stream is from an external source and is used to condense the steam. The reactor pressure vessel contains control rods and fuel rods (not labeled). It is within a larger containment structure. When water is heated by the reactor, it evaporates to steam. This moves through the steam line and turns a turbine, powering the electric generator. Steam cools in the condenser, and the cooled water is then pumped back into the reactor pressure vessel. Water from an external source is used to cool steam in the condenser. The water from the external source cools in the cone-shaped cooling tower. Image by Office of Nuclear Energy/ U.S. Department of Energy (public domain).

Attribution

Modified by Melissa Ha from the following sources:

- Non-renewable Energy Sources from *AP Environmental Science* by University of California College Prep (CC-BY). Download for free at CNX.
- Stages of the Nuclear Fuel Cycle. 2020. U.S. NRC. Accessed 01-16-2021 (public domain).
- Nuclear explained: The nuclear fuel cycle. 2020. U.S. Energy Information Administration. Accessed 01-16-2021 (public domain).

This page titled 6.2.2: Generating Electricity with Nuclear Energy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **17.2: Generating Electricity with Nuclear Energy** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

6.2.3: Nuclear Energy Consumption

Nuclear electricity came on the energy scene remarkably quickly. Following the development of nuclear technology at the end of World War II for military ends, nuclear energy quickly acquired a new peacetime path for inexpensive production of electricity. Eleven years after the end of World War II, a very short time in energy terms, the first commercial nuclear reactor produced electricity at Calder Hall in Sellafield, England. The number of nuclear reactors grew steadily to more than 400 by 1990 (figure 6.2.3.a), four years after the Chernobyl disaster in 1986 and eleven years following Three Mile Island in 1979 (see Consequences of Nuclear Energy). The number of operating reactors remained approximately flat for two decades, and the United States has not built a new nuclear facility since 1996. The number of operating nuclear power plants decreased in 2011, when the meltdown at the Fukushima Daiichi Nuclear Power Station caused Japan to shut down all of its nuclear power plants. Japan has since resumed use of some of its nuclear reactors.



Figure 6.2.3.a: Arkansas Nuclear One, a power plant in Russellville, Arkansas. Image by Edibobb (CC-BY).

World production of electricity from nuclear power was about 2795.96 terawatt-hours (TWh) in 2019, comprising 10.4% of electricity production and 4.3% of total energy consumption globally (figure 6.2.3.b). (For reference, the U.S. generated about 4100 TWh of electricity total in 2019.) The United States produced and consumed about 30.5% of the world's nuclear power in 2019, where nuclear power provided about 19.6% of the electricity and 8.0% of the total energy consumption (figure 6.2.3.c).

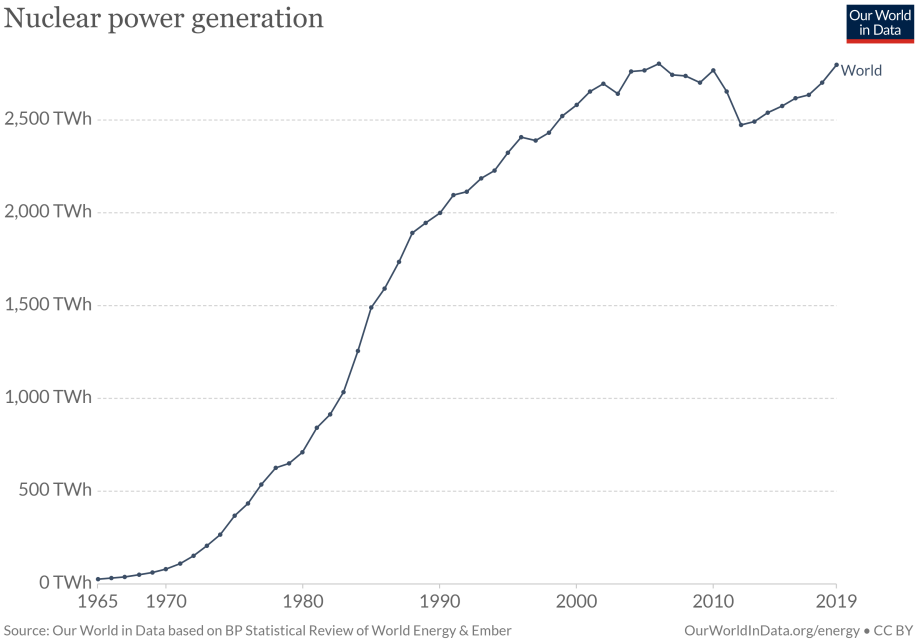


Figure 6.2.3.*b*: Global nuclear power generation in TWh over time. Nuclear power production initially increased but remained relative flat in the early 2000s. The dip in 2011 is due to the temporary shutdown of nuclear reactors in Japan following the Fukushima Daiichi Nuclear Power Station meltdown. Image from Hannah Ritchie/Our World in Data (CC-BY).

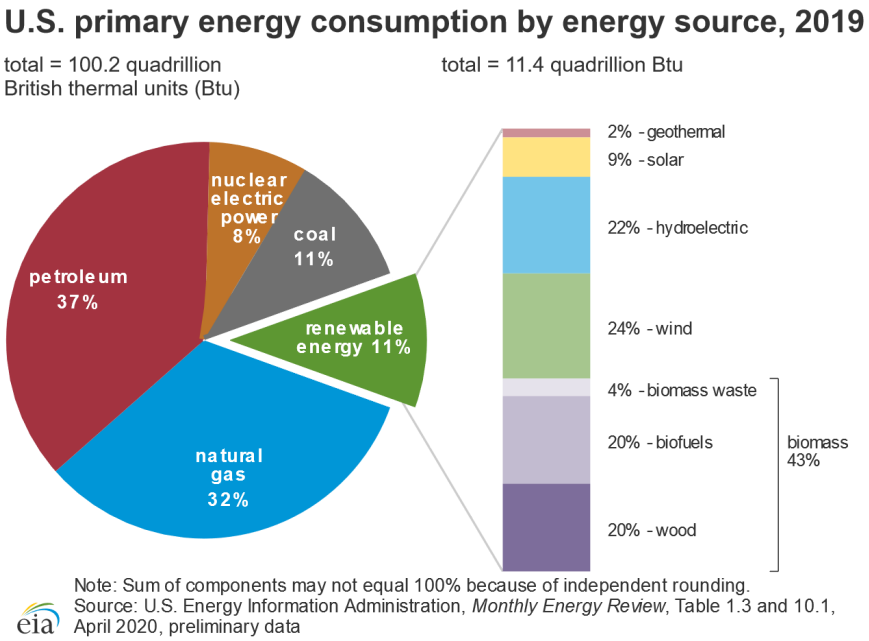


Figure 6.2.3.*c*: U.S. primary energy consumption by energy source in 2019. Nuclear power comprised 8% of total energy consumption in the U.S. Total energy consumption was 100.2 quadrillion British thermal units (Btu). Other source of energy were petroleum (37%), natural gas (32%), coal (11%), and renewable energy (11%).Total renewable energy consumption was 11.4 quadrillion Btu. This includes biomass (43%), wind (24%), hydroelectric (22%), solar (9%), and geothermal (2%). Biomass includes wood (20%), biofuels other than wood and waste (20%), and biomass waste (4%). Note: Sum of components may not equal 100% because of independent rounding. Image by EIA (public domain).

Attribution

Modified by Melissa Ha from Non-Renewable Energy Sources from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 6.2.3: Nuclear Energy Consumption is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **17.3: Nuclear Energy Consumption** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

6.2.4: Consequences of Nuclear Energy

The use of nuclear energy presents an interesting dilemma. On the one hand, nuclear electricity produces no carbon emissions, a major sustainable advantage in a world facing climate change. On the other hand, there is environmental risk of storing spent fuel for thousands or hundreds of thousands of years, societal risk of nuclear proliferation, and the impact of accidental releases of radiation from operating reactors. Thoughtful scientists, policy makers, and citizens must weigh these advantages and disadvantages.

Advantages of Nuclear Energy

In contrast to fossil fuels, generating electricity from nuclear energy does not pollute the air or significantly contribute to climate change (figure 6.2.4.*a*). As we continue to deplete global reserves of fossil fuels, supplies of nuclear fuel are abundant. It is estimated that uranium supplies will last over 200 years, and there is potential to use other radioactive isotopes as well. Furthermore, nuclear power plants are more reliable than any other source, with a capacity factor of 93.5% (figure 6.2.4.*b*). **Capacity** is the amount of electricity a generator can produce when it's running at full blast, and the **capacity factor** is a measure of how often a plant is running at maximum power. (A power plant with a capacity factor of 100% means that it's producing power all of the time.)

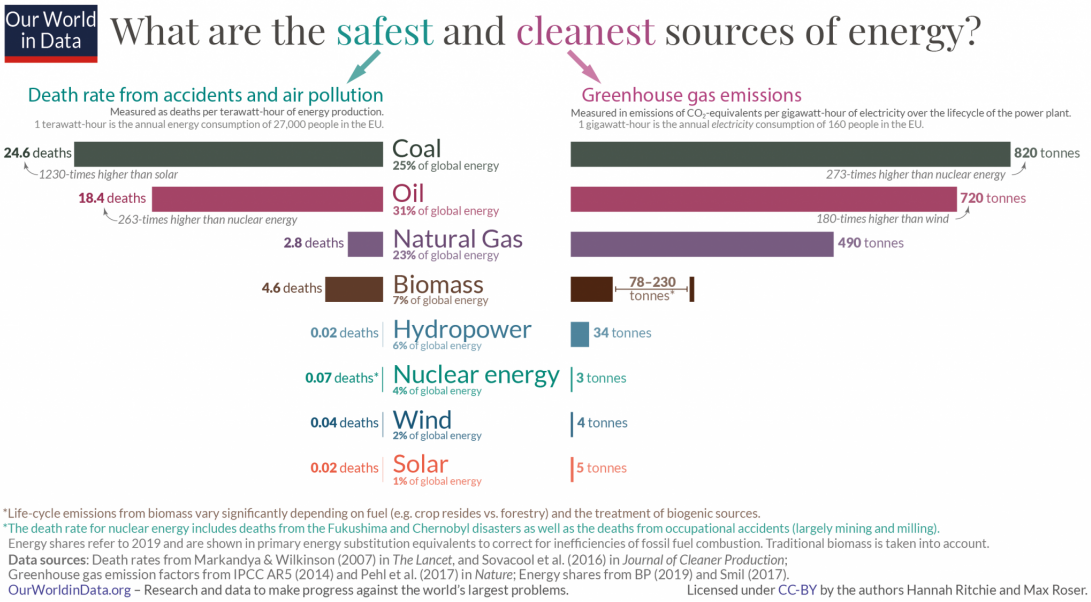


Figure 6.2.4.*a*: Nuclear energy causes fewer deaths and releases miniscule greenhouse gases compared to fossil fuels. These bar graphs explore the question, "What are the safest and cleanest sources of energy?". On the left is the death rate from accidents and air pollution measured as deaths per terawatt-hour (deaths/TWh) of energy production. 1 terawatt-hour is the annual energy consumption of 27,000 people in the European Union (EU). On the right are greenhouse gas emissions measured in tonnes of CO2-equivalents emitted per gigawatt-hour (tCO_2-eq/gWh) of electricity over the lifecycle of the power plant. 1 gigawatt-hour is the annual electricity consumption of 160 people in the EU. Coal accounted for 24.6 deaths/TWh, 25% of global energy, and 820 tCO_2-eq/gWh. Oil accounted for 18.4 deaths/TWh, 31% of global energy, and 720 tCO_2-eq/gWh. Coal accounted for 2.8 deaths/TWh, 23% of global energy, and 490 tCO_2-eq/gWh. Biomass accounted for 4.6 deaths/TWh, 7% of global energy, and 78-230 tCO_2-eq/gWh. Hydropower accounted for 0.02 deaths/TWh, 6% of global energy, and 34 tCO_2-eq/gWh. Nuclear energy accounted for 0.07 deaths/TWh, 4% of global energy, and 4 tCO_2-eq/gWh. Wind accounted for 0.04 deaths/TWh, 2% of global energy, and 4 tCO_2-eq/gWh. Coal accounted for 0.02 deaths/TWh, 1% of global energy, and 5 tCO_2-eq/gWh. Coal caused 1230 times more deaths than solar, and oil caused 263 times more deaths than nuclear energy. Coal emissions were 273 times higher than nuclear energy, and oil emissions were 180 times higher than wind. Image by Hannah Ritchie and Max Roser/Our World in Data (CC-BY).

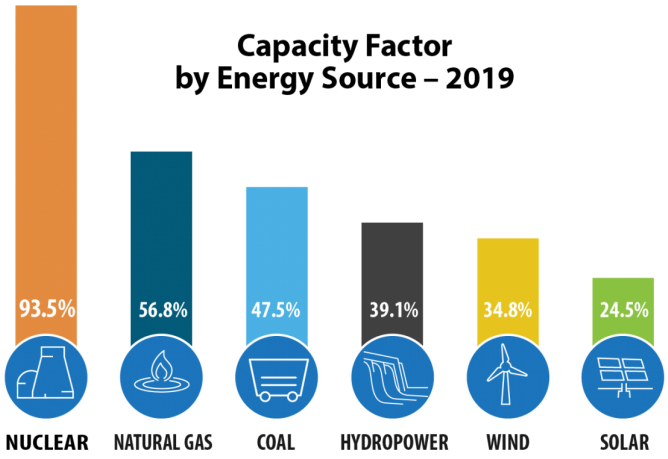


Figure 6.2.4.*b*: The capacity factor of six energy sources in 2019. Nuclear has the highest capacity factor at 93.5% followed by natural gas (56.8%), coal (47.5%), hydropower (39.1%), wind (34.8%), and solar (24.5%). Image from Office of Nuclear Energy/ U.S. Department of Energy (public domain).

Negative Impacts of Nuclear Energy

Despite its benefits, nuclear power has downsides. It requires more water than any other energy source. Water used for cooling is released back into the environment, and while it does not contain radioactive materials or other harmful chemicals, it is warmer than before. This is called **thermal pollution**, and it can harm aquatic life, which are adapted to cooler temperatures. Surface mining for uranium ore degrades habitat and releases toxins from underground (similar to surface mining for coal). Nuclear power plants are expensive to build and maintain, and they require large amounts of metal and concrete. Enriched uranium for nuclear fuel if in the wrong hands can be used to make nuclear weapons (figure 6.2.4.*c*). While nuclear accidents are rare, they can cause great harm, and their impacts are long-lasting. Furthermore, the problem of safely disposing spent nuclear fuel remains unresolved. The latter two concerns are discussed in more detail below.

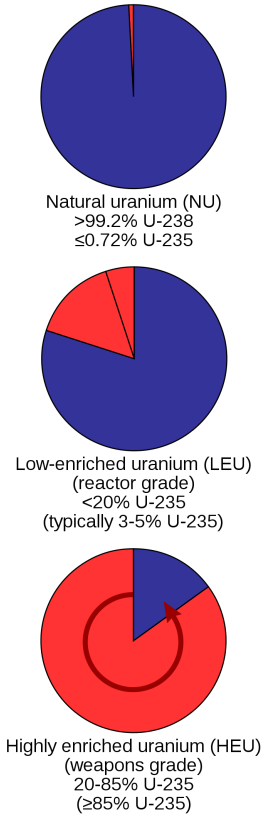


Figure 6.2.4.*c*: Pie charts showing the relative proportions of uranium-238 (blue) and uranium-235 (red) at different levels of enrichment. Nuclear weapons require more highly enriched uranium than is needed for nuclear fuel. Natural uranium (NU) consists of more than 99.2% U-238 and 0.72% or less of U-235. Low-enriched uranium (LEU) consists of less than 20% U-235, with reactor-grade uranium typically 2-5% U-235. High-enriched uranium (HEU) consists of 20-85% U-235 with weapons grade at least 85% U-235. Image and caption (modified) by Fastfission (public domain).

Nuclear Waste

The main environmental challenge for nuclear power is the wastes including high-level radioactive waste, low-level radioactive waste, and uranium mill tailings. These materials have long radioactive half-lives and thus remain a threat to human health for thousands of years.

High-level radioactive waste (HLRW) consists of used nuclear reactor fuel (spent nuclear fuel rods). These contain the products of nuclear fission, which are radioactive themselves. This HLRW is temporarily stored in a pool at the nuclear power plant or a dry cask, steel cylinders within another container, made of steel or concrete (figure 6.2.4.*d*). Dry casks contain inert (nonreactive) gas and may be located at the power plant, a decommissioned power plant, or a separate storage site. High-level radioactive waste may only be moved to a dry cask after one year of cooling in a pool. The U.S. has no long-term storage for HLRW, and spent fuel thus remains interim storage.

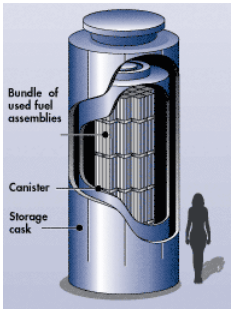


Figure 6.2.4.*d*: Dry casks seal spent fuel rods in a canister, which is surrounded by a larger storage cask. Bundles of used fuel assemblies are inside. Image by NRC (public domain).

Yucca Mountain in Nevada was proposed as a long-term geologic storage site, where HLRW could be buried for hundreds of thousands of years. The storage facility was constructed, but it has not been used due to opposition from local residents and concern about the safety of transporting HLNW (figure 6.2.4.*e*)



Figure 6.2.4.*e*: Yucca Mountain is the proposed site for long-term storage of high-level radioactive waste in the U.S., but it is not in use due to political controversy. Image by whitehouse.gov (public domain).

Some countries reprocess (recycle) spent nuclear fuel, but no recycling or reprocessing facility or a federal waste repository is currently licensed in the United States. Reprocessing separates the useable fraction of spent fuel and recycles it through the reactor, using a greater fraction of its energy content for electricity production, and sends the remaining high-level waste to permanent geologic storage.

The primary motivation for reprocessing is greater use of fuel resources, extracting about 25 percent more energy than the once through cycle. A secondary motivation for recycling is a significant reduction of the permanent geologic storage space (to 20% or less of what would otherwise be needed) and time (from hundreds of thousands of years to thousands of years). While these advantages seem natural and appealing from a sustainability perspective, they are complicated by the risk of theft of nuclear material from the reprocessing cycle for use in illicit weapons production or other non-sustainable ends. At present, France, the United Kingdom, Russia, Japan, and China engage in some form of reprocessing; the United States, Sweden, and Finland do not reprocess.

Low-level radioactive waste (LLRW) refers to items that were exposed to radiation includes clothing, filters, and gloves. These can be contained with concrete or lead (through which radiation cannot pass; figure 6.2.4.*f*). Low-level waste is typically stored at the nuclear power plant, either until it has decayed away and can be disposed of as ordinary trash, or until amounts are large enough for shipment to one of the five LLRW disposal sites in the U.S. (figure 6.2.4.*g*).

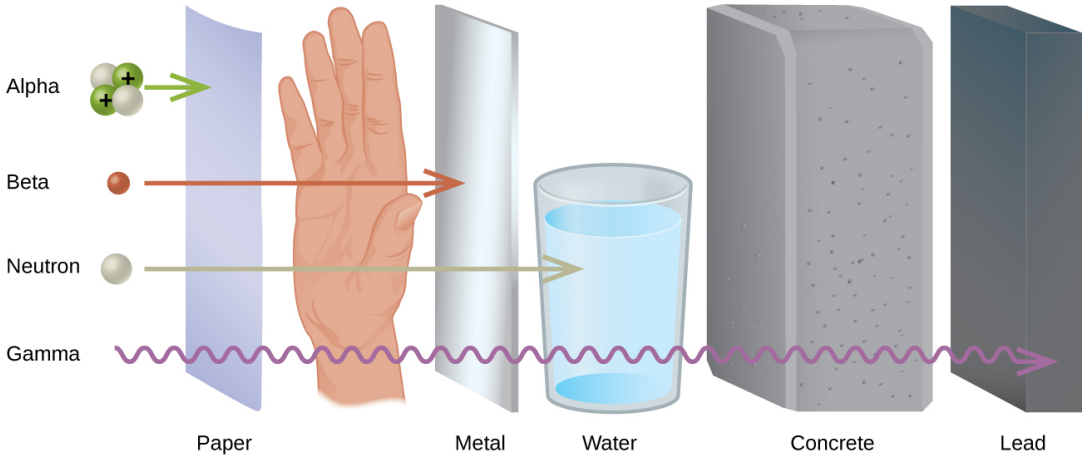


Figure 6.2.4.f: Water, thick concrete, lead, and steel (not shown) can stop several types of radiation released from radioactive wastes. Note that gamma rays can somewhat penetrate all of these substances, but lead, concrete, and steel provide a partial shield. The ability of different types of radiation to pass through material is shown. From least to most penetrating, they are alpha < beta < neutron < gamma. The top particle listed is made up of two white spheres and two green spheres that are labeled with positive signs and is labeled "Alpha." A right-facing arrow leads from this to the paper. The second particle is a red sphere labeled "Beta" and is followed by a right-facing arrow that passes through the paper and stops at the hand. The third particle is a white sphere labeled "Neutron" and is followed by a right-facing arrow that passes through the paper, hand and metal but is stopped at the glass of water. The fourth particle is shown by a squiggly arrow and it passes through all of the substances but stops at the lead. Terms at the bottom read, from left to right, "Paper," "Metal," "Water," "Concrete" and "Lead." Image and caption (modified) from Flowers, Theopold, and Langley/OpenStax (CC-BY). Download for free at CNX.

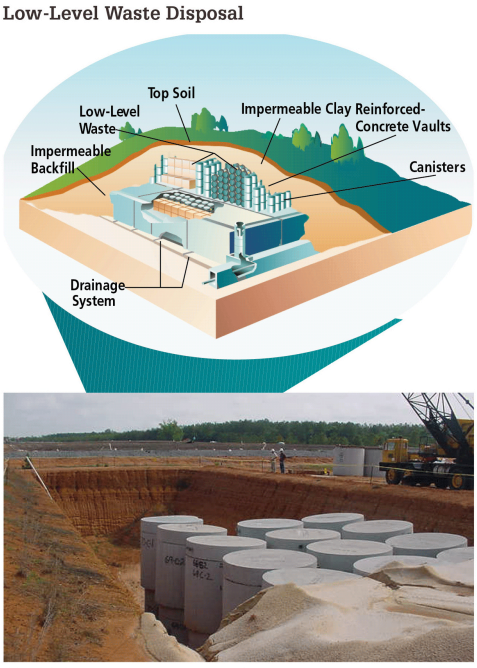


Figure 6.2.4.g: Diagram (top) and photo (bottom) of a low-level radioactive waste (LLRW or LLW) disposal site. In the diagram, low-level waste is contained in canisters within concrete vaults. This is surrounded by impermeable clay and backfill. The drainage system prevents the wastes from contaminating groundwater. The entire disposal facility is underground, and a layer of top soil is above it. The disposal site in the photo accepts waste from States participating in a regional disposal agreement.The waste is sealed in canisters and shallowly buried. Image by NRC (public domain).

Enrichment of uranium produces depleted uranium hexafluoride (DUF_6), or **uranium mill tailings**, as a byproduct, which does not have high enough concentrations of ^{235}U to use as nuclear fuel but is still hazardous. Tailings represent the greatest percentage of nuclear waste by volume, and there are more than 200 million tons of radioactive mill-tailings in the United States. Tailings contain

several radioactive elements including radium, which decays to produce radon, a radioactive gas. They are stored in **impoundments**, lined pits in the ground that are flooded with water, in remote areas. **Deconversion** involves chemically treating the tailings to reduce their hazards so that they can be stored as LLRW.

Nuclear Disasters

There are many other regulatory precautions governing permitting, construction, operation, and decommissioning of nuclear power plants due to risks from an uncontrolled nuclear reaction. The potential for contamination of air, water and food is high should an uncontrolled reaction occur. Even when planning for worst-case scenarios, there are always risks of unexpected events. The nuclear accidents at Three Mile Island, Chernobyl (see the Chapter Hook), and Fukushima raised concerns about the safety of nuclear power.

The Three Mile Island accident occurred in Pennsylvania in 1979. It was a partial meltdown that resulted from an electrical failure and errors in operation. There were no direct deaths. Studies investigated the possibility of exposure to radiation from the accident indirectly causing deaths through increased rates of cancer or other disease, but there has not been evidence of this. In contrast, the 1986 meltdown at Chernobyl Nuclear Power Plant in what is now the Ukraine was responsible for 50 direct deaths. This disaster occurred from a test of the emergency systems gone wrong. Estimates of indirect deaths from radiation exposure range from 4,000 to 60,000.

The global discussion regarding nuclear energy has been strongly impacted by March 2011 earthquake and subsequent tsunami that hit Japan resulted in reactor meltdowns at the Fukushima Daiichi Nuclear Power Station causing massive damage to the surrounding area. The disaster disabled the cooling system for a nuclear energy complex, ultimately causing a partial meltdown of some of the reactor cores and release of significant radiation. The design of the reactors (boiling water reactors) made it more difficult to vent the system without releasing radiation. Cooling the radioactive fuel generated a large volume of contaminated water, and the disaster costed at least $300 billion dollars. While there were no immediate deaths, one person later died from cancer attributed to radiation exposure. Thousands died as a result of stress associated with the evacuation, and about 20% of the over 160,000 evacuees had not yet returned home as of 2019.

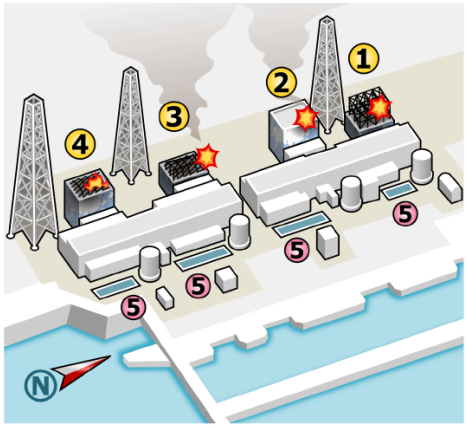




Figure 6.2.4.*g*: Left: Diagram (approximate) of the Fukushima I Nuclear Power plant accidents. (1) Unit 1: Explosion, roof blown off on 12 March. (2) Unit 2: Explosion on 15 March; contaminated water in underground trench, possible leak from suppression chamber. (3) Unit 3: Explosion, most of concrete building destroyed on 14 March, possible plutonium leak. (4) Fire on 15 March; water level in spent fuel pools partly restored. (5) Multiple trenches: probable source of contaminated water, partly underground, leaked stopped on 6 April. Right: The Fukushima I Nuclear Power Plant after the 2011 Tōhoku earthquake and tsunami. Reactor 1 to 4 from right to left. Left image and caption (modified) by Sodacan (CC-BY). Right image and caption by Digital Globe (CC-BY-SA).

Attribution

Modified by Melissa Ha from the following sources:

- Non-Renewable Energy Sources from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Nuclear Energy, Environmental Challenges in Energy, Carbon Dioxide, Air, Water and Land Use, and Systems of Waste Management from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors. Textbook content produced by Tom Theis and Jonathan (CC-BY). Download for free at CNX.
- Non-renewable Energy Sources from *AP Environmental Science* by University of California College Prep (CC-BY). Download for free at CNX.
- Stages of the Nuclear Fuel Cycle and Low-Level Waste. 2020. U.S. NRC. Accessed 01-16-2021 (public domain).
- What is Generation Capacity? 2020. Mike Mueller. Office of Nuclear Energy. U.S. Department of Energy. Accessed 01-16-2021 (public domain)

This page titled 6.2.4: Consequences of Nuclear Energy is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **17.4: Consequences of Nuclear Energy** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

6.2.5: Data Dive- Global Nuclear Power Generation

Overview

Our World in Data (OWID) is a scientific online publication that focuses on using research and data to help tackle global concerns such as poverty, disease, hunger, climate change, war, and inequitable treatment of our world's most vulnerable and unstable communities. Their website is particularly known for publishing a variety of graphs, some even interactive, to present the research that helps explain causes and consequences of global concerns to the public. One example graph, seen below, illustrates nuclear power generation:

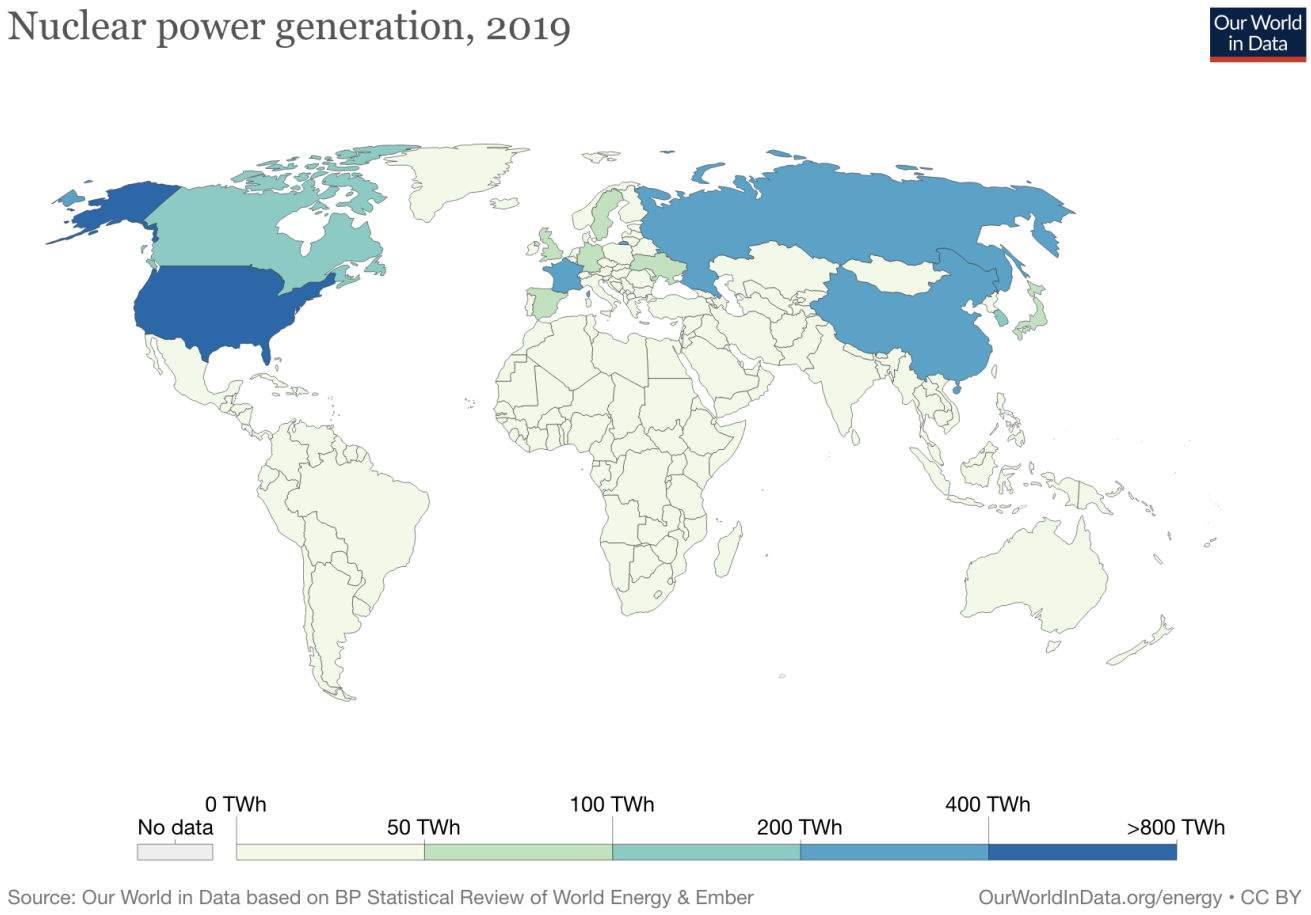


Figure 6.2.5.*a*: Global nuclear power generation in 2019. Graph by Our World in Data (CC-BY)

Questions

1. What countries generate the most energy via nuclear power?
2. Is there anything that surprises you about the patterns observed on this graph?
3. Based on what you learned in this chapter about nuclear power, do you think the trends you see in this 2019 map continue for the next 50 years? Explain.
4. Again, what is missing on this map?

Raw Data For Above Graph(s)

Raw data table link for Our World In Data graph above. (CC-BY)

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 6.2.5: Data Dive- Global Nuclear Power Generation is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **17.5: Data Dive- Global Nuclear Power Generation** by Melissa Ha and Rachel Schleiger has no license indicated.

6.2.6: Review

Summary

After completing this chapter you should be able to...

- Define isotopes and radioactive isotopes.
- Explain how half-life measure the rate of radioactive decay.
- Explain how nuclear fission reactions can be induced to generate electricity.
- Outline the nuclear fuel cycle.
- Describe the structure and functioning of a nuclear reactor, distinguishing between pressurized water reactors and boiling water reactors.
- Detail the percentage that nuclear power contributes to electricity and total energy consumption globally and in the United States.
- Discuss the advantages and disadvantages of nuclear energy.

Isotopes are atoms of the same element that differ in neutron number. Some isotopes are **radioactive**, meaning that they are unstable and emit radiation in the form of particles and energy. The speed of this radioactive decay is measured in **half-lives**. A **nuclear fission reaction** is the splitting of an atom nucleus. Nuclear fission of uranium-235 can be induced by neutrons, and this is the basis of nuclear power.

The **nuclear fuel cycle** describes the process of mining uranium ore, milling it into yellowcake, and enriching it to produce nuclear fuel. It also outlines proper storage and disposal of spent fuel and other waste products.

Nuclear fission occurs in the **reactor core** of a **nuclear reactor**. It is a chain reaction in which fission releases additional neutrons that induce fission in other atoms. **Fuel rods** contain nuclear fuel while **control rods** absorb excess neutrons to prevent an explosion or meltdown. The heat released from nuclear fission ultimately generates high-pressure steam, which turns a turbine, powering a generator. In this sense the process of nuclear electricity generation is similar to that from coal. **Pressurized water reactors** consist of three separate streams of water, but **boiling water reactors** consist of two separate streams of water.

Nuclear power accounts for 10.4% of electricity production and 4.3% of total energy consumption globally. In the United States, it accounts for 9.6% of the electricity and 8.0% of the total energy consumption.

Nuclear energy is beneficial in the sense that it releases few greenhouse gases or air pollutants. It is a reliable energy source, and nuclear fuel is abundant. However, nuclear power plants consume more water than any other source of energy. Power plants are expensive to build and maintain. While there have not been many nuclear accidents globally, some have been deadly, and residents still suffer from their effects. Additionally, nuclear waste continues to emit dangerous radiation, and the U.S. does not have a long-term storage facility for high-level radioactive waste.

Attribution

Melissa Ha (CC-BY-NC)

This page titled 6.2.6: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **17.6: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

6.3: Renewable Energy

Chapter Hook

City water pipes, an unlikely place to be a source electricity. Yet, Portland Oregon has decided to set up hydroelectric turbines in the cities drinking water pipes. This provides around the clock local energy that doesn't affect natural waterways as classic hydroelectric turbines do. This setup is the first of its kind in the United States and truly exemplifies the objectives that renewable energies aspire to achieve.



Figure 6.3.*a*: Water turbine example. Image by Pixabay (Public Domain)

6.3.1: Renewable Energy History and Consumption

6.3.2: Wind Energy

6.3.3: Solar Energy

6.3.4: Geothermal Energy

6.3.5: Hydropower

6.3.6: Biofuels (Biomass Energy)

6.3.7: Energy Conservation

6.3.8: Data Dive- Global Renewable Energy Generation

6.3.9: Review

Attribution

Modified by Rachel Schleiger (CC-BY-NC).

This page titled 6.3: Renewable Energy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

6.3.1: Renewable Energy History and Consumption

Renewable energy sources can be replenished within human lifespans. Although renewable energy is often classified as **wind**, **solar**, **geothermal**, **hydropower** (hydroelectric energy/hydroelectricity), and **biofuels** (**biomass energy**), all forms of renewable energy arise from only three sources: the light of the sun (wind, solar, hydropower, and biofuels), the heat of the earth's crust (geothermal), and the gravitational attraction of the moon and sun (**tidal energy**; figure 6.3.1.a). Sunlight provides by far the largest contribution to renewable energy. The sun provides the heat that drives the weather, including the formation of high- and low-pressure areas in the atmosphere that make wind. The sun also generates the heat required for vaporization of ocean water that ultimately falls over land creating rivers that drive hydropower, and the sun is the energy source for photosynthesis, which creates biomass. The sun is also responsible for the energy of fossil fuels, created from the organic remains of plants and sea organisms compressed and heated in the absence of oxygen in the earth's crust for tens to hundreds of millions of years. The time scale for fossil fuel regeneration, however, is too long to consider them renewable in human terms.



Figure 6.3.1.a: A tidal current energy generator at Race Rocks in British Columbia, Canada. Image by Garry Fletcher (CC-BY-SA).

Most renewable energy sources have a relatively **small carbon footprint**, meaning that they do not significantly contribute to climate change. However, building dams for hydropower (hydroelectric energy) does release methane, a potent greenhouse gas. Renewable energy sources are pollution free (except biofuels), and they typically have minimal environmental impact (except habitat loss from dams). Up to this point, however, no single source of renewable energy is sufficient. They generally are paired with other energy sources.

Strong interest in renewable energy in the modern era arose in response to the oil shocks of the 1970s, when the Organization of Petroleum Exporting Countries (OPEC) imposed oil embargos and raised prices in pursuit of geopolitical objectives. The shortages of oil, especially gasoline for transportation, and the eventual rise in the price of oil by a factor of approximately 10 from 1973 to 1981 disrupted the social and economic operation of many developed countries and emphasized their precarious dependence on foreign energy supplies. The reaction in the United States was a shift away from oil and gas to plentiful domestic coal for electricity production and the imposition of fuel economy standards for vehicles to reduce consumption of oil for transportation. Other developed countries without large fossil reserves, such as France and Japan, chose to emphasize nuclear (France to the 80% level and Japan to 30%) or to develop domestic renewable resources such as hydropower and wind (Scandinavia), geothermal (Iceland), solar, biomass and for electricity and heat. As oil prices collapsed in the late 1980s, interest in renewables, such as wind and solar that faced significant technical and cost barriers, declined in many countries. Other renewables, such as hydropower and biomass, continued to experience growth. As climate change worstens, the solar and wind energy has expanded globally in recent years.

Renewable energy accounted for 11.4% of total global energy consumption and 26.3% of global electricity generation in 2019 (figure 6.3.1.*b*). In the United States, renewable energy also accounted for about 11% of total energy consumption but only 17.6% of electricity generation. Nearly half (43%) of total renewable energy use in the U.S. is from biofuels. For electricity generation in the U.S., wind was the greatest contributor (7.1%) followed by hydropower (7.0%), solar (1.7%), biomass (1.4%), and geothermal (0.4%).

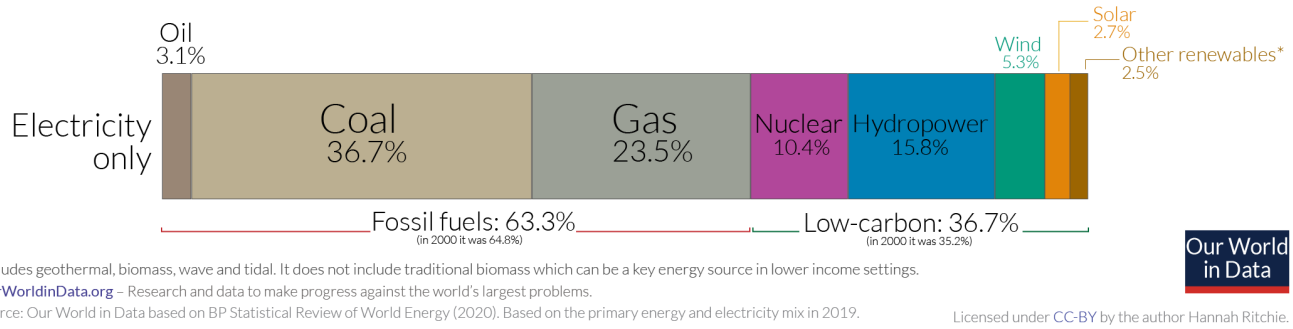


Figure 6.3.1.*b*: The global electricity breakdown in 2019. Renewables accounted for 26.3% of global electricity generation, with hydropower accounting for 15.8% followed by wind (5.3%), solar (2.7%), and other renewables (2.5%; geothermal, biomass, wave, and tidal). Fossil fuels accounted for 63.3% of global electricity generation, including oil (3.1%), coal (36.7%), and gas (23.5%). For comparison, fossil fuels accounted for 64.8% of global electricity generation in 2000. Nuclear accounted for 10.4% of global electricity generation. Nuclear and renewables combined are considered low-carbon electricity sources, together accounting for 36.7% of electricity generation. For comparison, low-carbon sources accounted for 35.2% of electricity generation in 2000. Image modified from Hannah Ritchie/Our World in Data (CC-BY).

Attribution

Modified by Melissa Ha from Renewable Energy and Challenges and Impacts of Energy Use from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 6.3.1: Renewable Energy History and Consumption is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **18.1: Renewable Energy History and Consumption** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

6.3.2: Wind Energy

Wind energy arises from the motion of the air. It is driven by solar energy (differences in air temperature causes air currents). The wind turns a turbine, which powers a generator (figure 6.3.2.a). The rotor blades of a wind turbine work like an airplane wing or helicopter rotor blade. When wind flows across the blade, the air pressure on one side of the blade decreases, and this causes the rotor to spin. The rotor connects to the generator, either directly or through a series of gears that speed up the rotation and allow for a physically smaller generator. Similar to the electricity generation from coal, natural gas, or nuclear energy, the rotating motion causes magnets spin within wire coils to produce electricity. According to the American Wind Energy Association, 39% of all new electrical generating capacity in the United States in 2019 was due to wind.

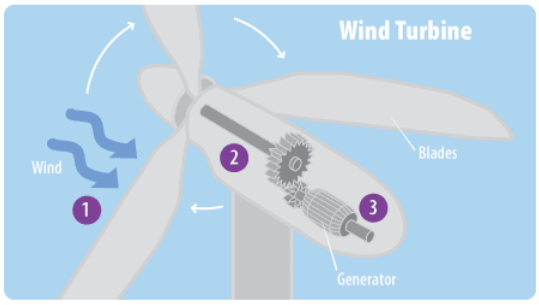


Figure 6.3.2.a: (1) As the wind blows over the blades of a wind turbine, it causes the blades to lift and rotate. (2) The rotating blades turn a shaft that is connected to a generator. (3) The generator creates electricity as it turns. Image and caption (modified) from EPA (public domain).

Advantages of Wind Energy

Wind is among the lowest-cost sources of renewable energy, and its expansion creates jobs (figure 6.3.2.b). Like many renewable energy sources, wind turbines do not release air pollutants or contribute to climate change, and they do not require water for cooling. Because a wind turbine has a small physical footprint relative to the amount of electricity it produces, many wind farms are located on crop and pasture land. They contribute to economic sustainability by providing extra income to farmers and ranchers, allowing them to stay in business and keep their property from being developed for other uses. For example, energy can be produced by installing wind turbines in the Appalachian mountains of the United States instead of engaging in mountain top removal for coal mining. Offshore wind turbines on lakes or the ocean may have smaller environmental impacts than turbines on land, and winds are up to 50% stronger and steadier offshore than on land (figure 6.3.2.c).

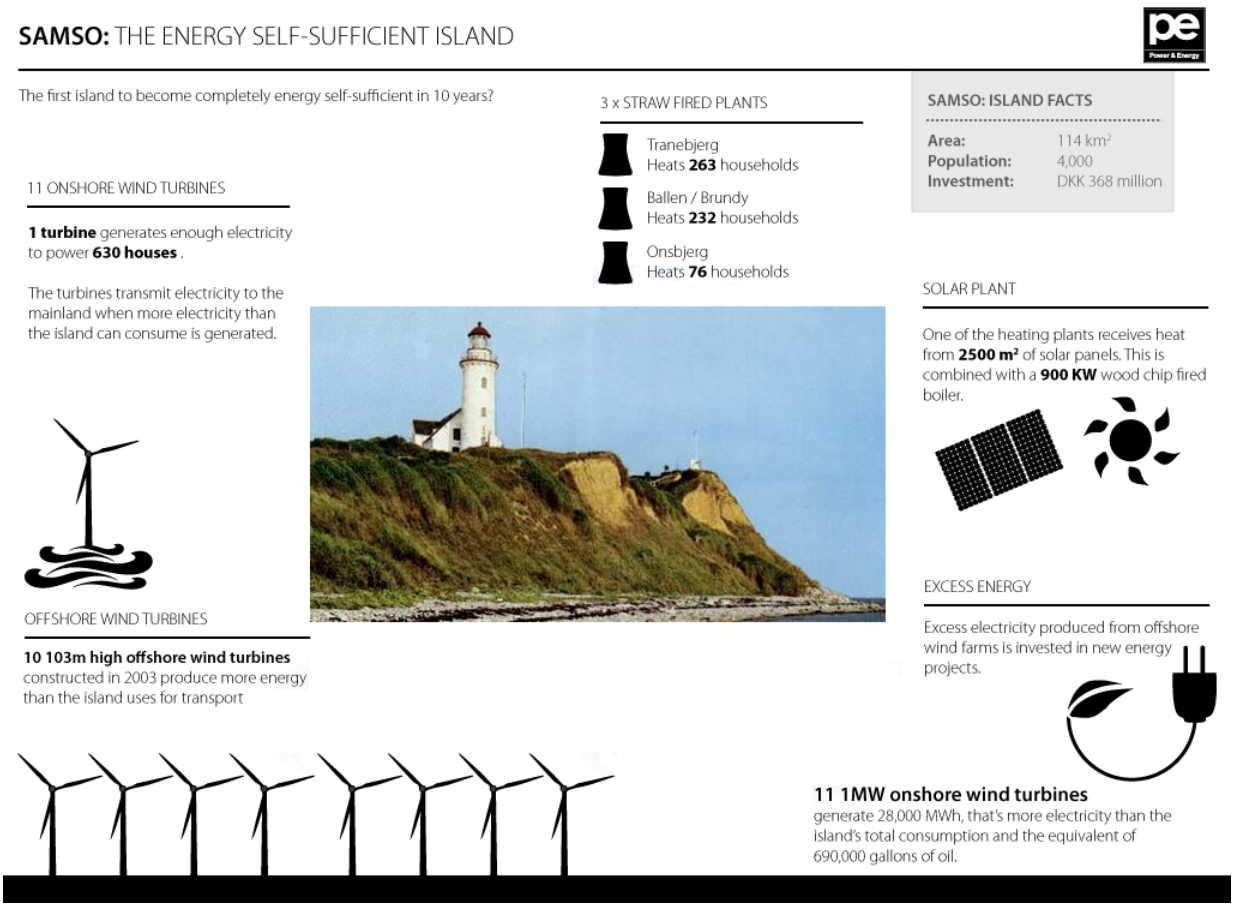


Figure 6.3.2.*b*: The Danish island of Samsø was the first region in the country to transition to 100% renewable energy, relying on offshore wind, onshore wind, straw (a biofuel), and solar. The text in the image is, "Samsø : The Energy Self-sufficient Island. The first island to become complete energy self-sufficient in 10 years. Samsø: Island Facts— Area: 114 km^2, Population: 4,000, Investment: DKK 368 million. 11 Onshore Wind Turbines— 1 turbine generates enough electricity to power 630 houses. The turbines transmit electricity to the mainland when more electricity than the island can consume is generated. Offshore Wind Turbines- 10 103-m-high offshore wind turbines constructed in 2003 produce more energy than the island uses for transport. 11 1-MW Onshore Wind Turbines generate 28,000 MWh. That's more electricity than the island's total consumption and the equivalent of 690,000 gallons of oil. 3X Straw-fired Plants— Traneberg- heats 263 households, Ballen/Brundy- heats 232 households, Onsbjerg- heats 76 households. Solar Plant- One of the heating plants receives heat from 2500 m^2 of solar panels. This is combined with a 900 KW wood chip-fired boiler. Excess Energy- Excess electricity produced from offshore wind farms is invested in new energy projects. Images by GDS Infographics (CC-BY) and National Geospatial-Intelligence Agency (public domain).



Figure 6.3.2.*b*: Wind turbines may be on land (onshore, left) or in the ocean (offshore, right). Left image by Joshua Winchell/USFWS (public domain). Right image by CGP Grey (CC-BY).

Disadvantages of Wind Energy

Wind energy does have a few challenges. Wind turbines are only effective in regions with strong enough winds to generate sufficient electricity. Even in regions with strong winds, wind availability is intermittent. This can be mitigated by the use batteries to store energy, but battery capacity, despite continuous technological advances, is still limited. There are aesthetic concerns to some people when they see them on the landscape, and some people do not like the sound that wind turbine blades make. A few wind turbines have caught on fire, and some have leaked lubricating fluids, though this is relatively rare. Turbines have been found to cause bird and bat deaths particularly if they are located along their migratory path, although communication towers and domestic cats are bigger threats. There are some small impacts from the construction of wind projects or farms, such as the construction of service roads, the production of the turbines themselves, and the concrete for the foundations.

Attribution

Modified by Melissa Ha from the following sources:

- Renewable Energy and Challenges and Impacts of Energy Use from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- The Power of Wind. Office of Energy Efficiency and Renewable Energy. U.S. Department of Energy. Accessed 01-17-2021 (public domain).

This page titled 6.3.2: Wind Energy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **18.2: Wind Energy** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

6.3.3: Solar Energy

Solar energy refers to heat or light energy from the sun. Solar energy is by far the most plentiful type of renewable energy, delivered to the surface of the Earth at a rate of 120,000 Terawatts (TW) per hour, compared to the global human use of 19.8 TW in the entire year of 2019. To put this in perspective, covering 1.2% of the Sahara desert with solar panels could meet Earths energy needs. Of course, this does not consider limitations on storage capacity and the ability to distribute that energy.

Technologies to harness solar energy may be passive or active. **Passive solar technologies** do not require complex equipment and can be as simple as using natural light from a window or skylight to illuminate a room (figure 6.3.3.a). Similarly, solar tubes are lined with reflective material and can concentrate light energy to better illuminate a room (figure 6.3.3.a). These are embedded in the ceiling as regular light fixtures would be.

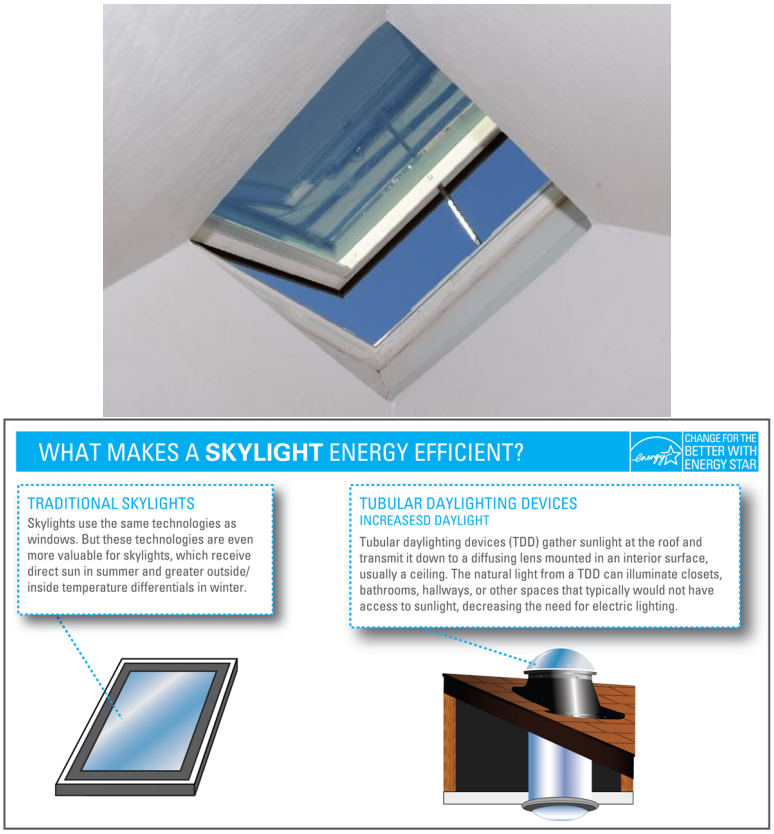


Figure 6.3.3.a: Skylights and solar tubes (tubular daylighting devices, TDD). While skylights resemble windows that are installed in the ceiling, solar tubes use lenses and a reflective lining to concentrate sunlight, providing bright illumination for darker parts of the home. What makes a skylight energy efficient? Traditional skylights are skylights that use the same technologies as windows, but these technologies are even more valuable for skylights, which receive direct sun in summer and greater outside/inside temperature differentials in winter. Tubular daylighting devices gather sunlight at the roof and transmit it down to a diffusing lens mounted in an interior surface, usually a ceiling. The natural light from a TDD can illuminate closets, bathrooms, hallways, or other spaces that typically would not have access to sunlight, decreasing the need for electric lighting. Left image by U.S. Department of Energy (public domain), and right image by Energy Star (public domain).

Solar energy can also be used as heat, which can be maximized through careful architecture (figure 6.3.3.b). Firstly, the building requires south-facing windows (or glass doors). As sunlight passes through these areas, the energy is stored in the thermal mass of the building. This refers to heat trapping materials such as rock or tiles. The building is also designed such that heat is then distributed throughout the building. Finally, roof eaves or similar structures block sunlight from entering the home during the summer.

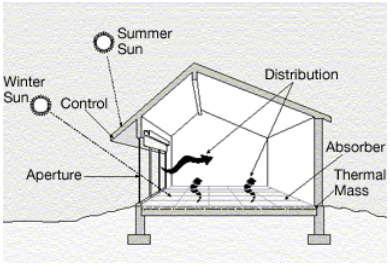


Figure 6.3.3.*b*: This building design uses the sun for heating during the winter and is thus a passive solar technology. The roof overhang serves as a control, blocking the higher summer sun while allowing the lower winter sunlight to pass through the window (aperture). Heat energy from the sun is then distributed in the building. Thermal mass (such as a tile or stone floor) absorbs and later releases heat energy. Image by U.S. Department of Energy (public domain).

A simple solar water heater is a passive technology consisting of a network of tubes that are heated by the sun (figure 6.3.3.*c*). The hot water is then transferred through the plumping of a home. (Some solar water heaters are more complex, using pumps, and are thus considered active solar technologies.)



Figure 6.3.3.*c*: A solar water heater. Image by Vijayanarasimha/Pixabay (public domain).

Active solar technologies are more complex. For example, solar panels use light energy to generate electricity (figure6.3.3.*d*). This occurs in the units of the solar panel, which are called **photovoltaic cells** (PV cells; figure6.3.3.*e*). Each photovoltaic cell consists of two layers of **semiconductors**, substances that only conduct electricity under certain circumstances. (In contrast, conductors always conduct electricity, and insulators do not.) One semiconductor has extra electrons, but the other has extra spaces for electrons. When light shines on the photovoltaic cell, it causes electrons to move from between semiconductor layers through the conductor that connects them (such as metal wires or plates). This movement results in an electric current.



Figure 6.3.3.*d*: Rooftop solar installation on Douglas Hall at the University of Illinois at Chicago has no effect on land resources, while producing electricity with zero emissions. Source: Office of Sustainability, UIC

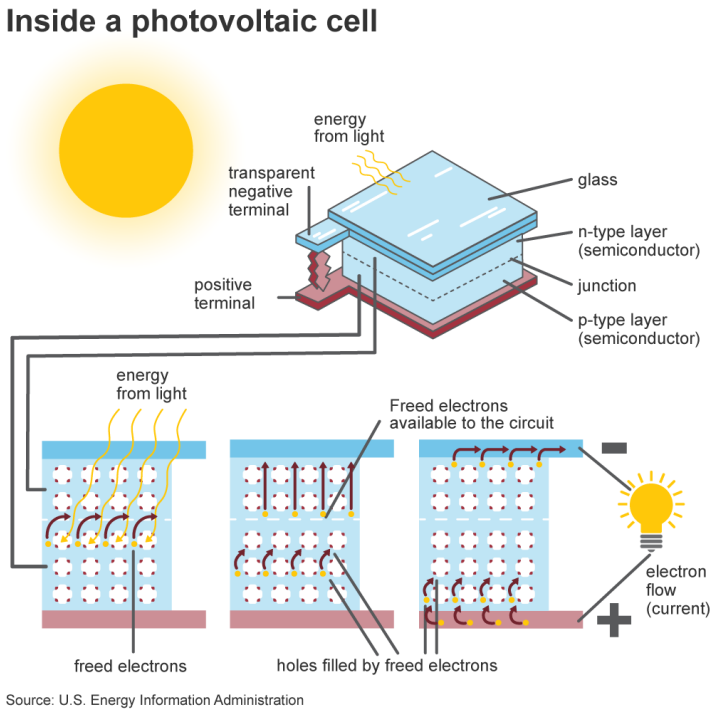
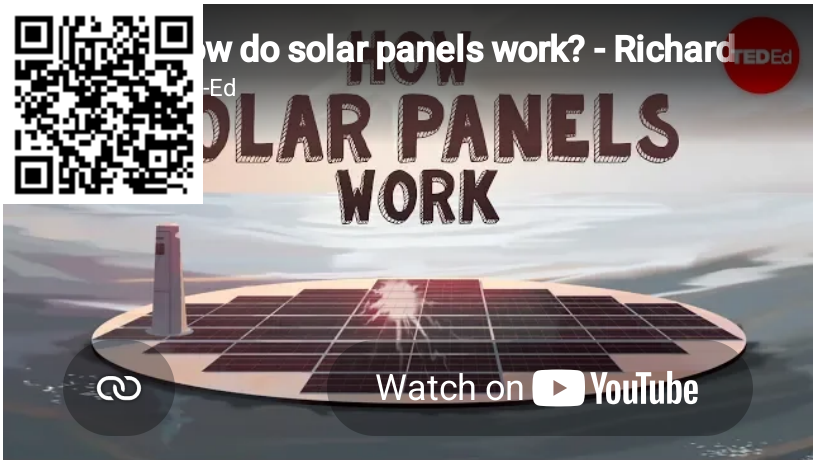


Figure 6.3.3.*e*: Diagram of a **photovoltaic cell** consisting of two layers of semiconductors. The top (n-type) semiconductor has extra electrons and the bottom (p-type) semiconductor has extra spots (holes) for electrons. A junction divides these two layers. Light frees the electrons and allows them to move through a wire from the n-type to fill "electron holes" in the p-type semiconductor. The movement of electrons through the wire results in an electrical current. Image by U.S. Energy Information Administration (public domain).

This video explains how the photovoltaic cells within solar panels generate electricity.



Another example of active solar technology is **solar thermal technology**. This involves using a series of mirror to concentrate solar energy, ultimately generating steam. From there, the steam turns a turbine and powers a generator (figure6.3.3.*f*).

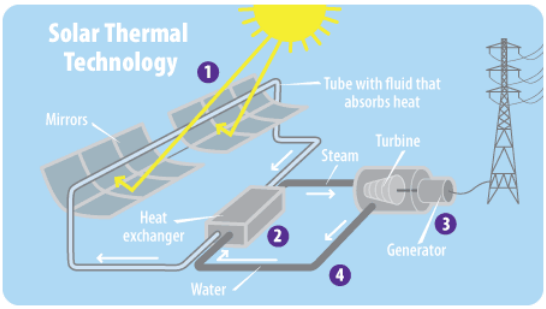


Figure 6.3.3.*f*: In solar thermal technology, (1) mirrors or reflectors concentrate the sun's rays to heat a special kind of liquid, which absorbs heat. (2) The heat from this liquid boils water to create steam, facilitated by the heat exchanger. (3) Steam spins a turbine that is connected to a generator, which creates electricity. (4) The steam cools and condenses back to water, which is recycled, reheated, and converted into steam again. Image and caption (modified) from EPA (public domain).

Not only is solar energy abundant, but the use of solar panels for electricity does not generate air pollution or contribute to climate change. (The manufacture of solar panels can generate some pollution, including greenhouse gas emissions, but this is minimal compared to that of fossil fuels.) Like wind energy, expansions on solar energy can create jobs and boost economies. Also like wind, sunlight is intermittent and storage of solar energy is limited by battery capacity. Some locations do not receive consistently direct sunlight and are ill-suited for solar panels. While solar energy has historically been the most expensive form of renewable energy, new technologies have lowered its cost.

The placement of solar panels determines how their environmental impact. Solar arrays are often placed on roofs of buildings or over parking lots or integrated into construction in other ways. However, large systems may be placed on land and particularly in deserts where those fragile ecosystems could be damaged if care is not taken. Additionally, solar farms can compete for agricultural space.

Other downsides of solar energy are water consumption (for some uses) and generation of hazardous wastes. Large networks of mirrors and lenses that concentrate solar energy for electricity generation in thermal solar systems or for heating may need to be cleaned regularly with water. Water is also needed for cooling the turbine-generator. Using water from underground wells may affect the ecosystem in some arid locations. The manufacturing of photovoltaic cells generates some hazardous waste from the chemicals and solvents used in processing. Some solar thermal systems use potentially hazardous fluids (to transfer heat) that require proper handling and disposal. Nuclear power exceeds solar energy in water consumption and hazardous waste generation, however.

Attribution

Modified by Melissa Ha from the following sources:

- Renewable Energy and Challenges and Impacts of Energy Use from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 6.3.3: Solar Energy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **19.3: Solar Energy** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

6.3.4: Geothermal Energy

Geothermal energy originates from heat rising to the surface from Earth's molten iron core created during the formation and compression of the early Earth as well as from heat produced continuously by radioactive decay of uranium, thorium, and potassium in the Earth's crust. **Geothermal power plants** harness this heat energy to produce electricity much in the same way that heat from burning coal generates energy (figure $6.3.4.a-c$). Water is injected underground and heated. The steam that emerges can be used directly, the heat can be transferred to closed system of another fluid, which then boils (figure $6.3.4.c$). Either way, the steam (or other high-pressure gas) ultimately turns a turbine and powers a generator.



Figure $6.3.4.a$: Steam is released from a geothermal power plant. Image by Open Access Government (CC-BY).

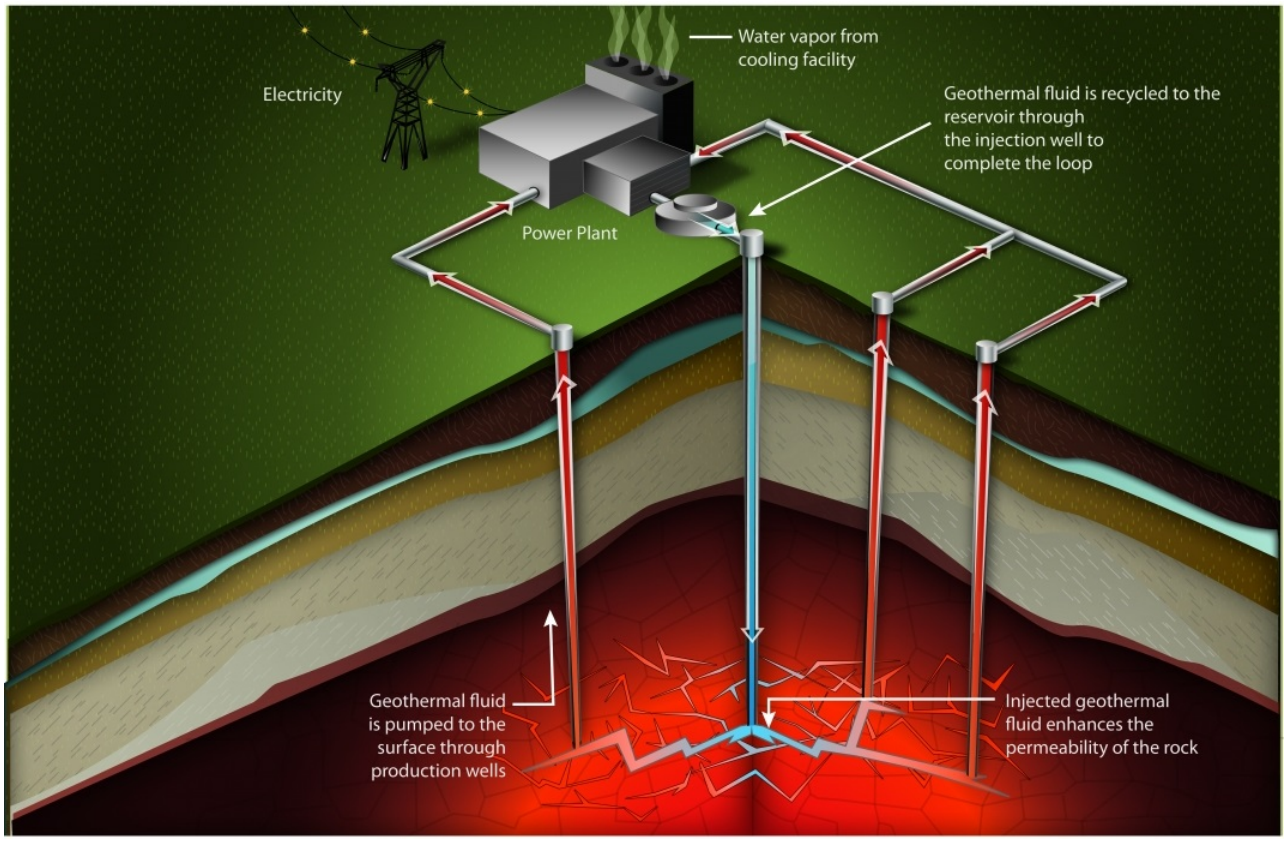


Figure 6.3.4.*b*: In a geothermal power plant, geothermal fluid is injected underground. (Geothermal water refers to the hot water/steam used in this process.) It enhances the permeability of the rocks and is heated underground. It exits to the surface through production wells, where it is used to generate electricity in the geothermal power plant. Water vapor is released from the cooling facility of the power plant. To recycle the geothermal fluid, it is returned to the reservoir and injected back into the ground. Image from Office of Energy Efficiency and Renewable Energy/U.S. Department of Energy (public domain).

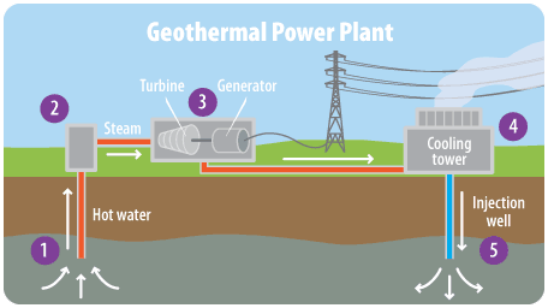


Figure 6.3.4.*c*: (1) In a geothermal power plant, hot water is pumped from deep underground through a well under high pressure. (2) When the water reaches the surface, the pressure is dropped, which causes the water to turn into steam. (3) The steam spins a turbine, which is connected to a generator that produces electricity. (4) The steam cools off in a cooling tower and condenses back to water. (5) The cooled water is pumped back into the Earth to begin the process again. Image and caption (modified) from EPA (public domain).

Geothermal heat pumps (**ground-source heat pumps**) rely on cool temperatures underground to cool or heat homes (figure 6.3.4.*d*). They are sometimes considered a second type of geothermal energy, but they are also a means of energy conservation. Geothermal heat pumps use a heat-exchange system that runs in the subsurface about 20 feet (5 meters) below the surface, which is consistently cool (around 55°F, or 12.5° C). Fluid is pumped underground and then along ducts in the home. This cools the house during the summer, acting as a heat sink. During a cold winter, it warms the house to 55° F (acting as a heat source), and traditional heating systems do the rest. This reduces the energy consumption required to generate heat from gas, steam, hot water, and conventional electric air-conditioning systems.

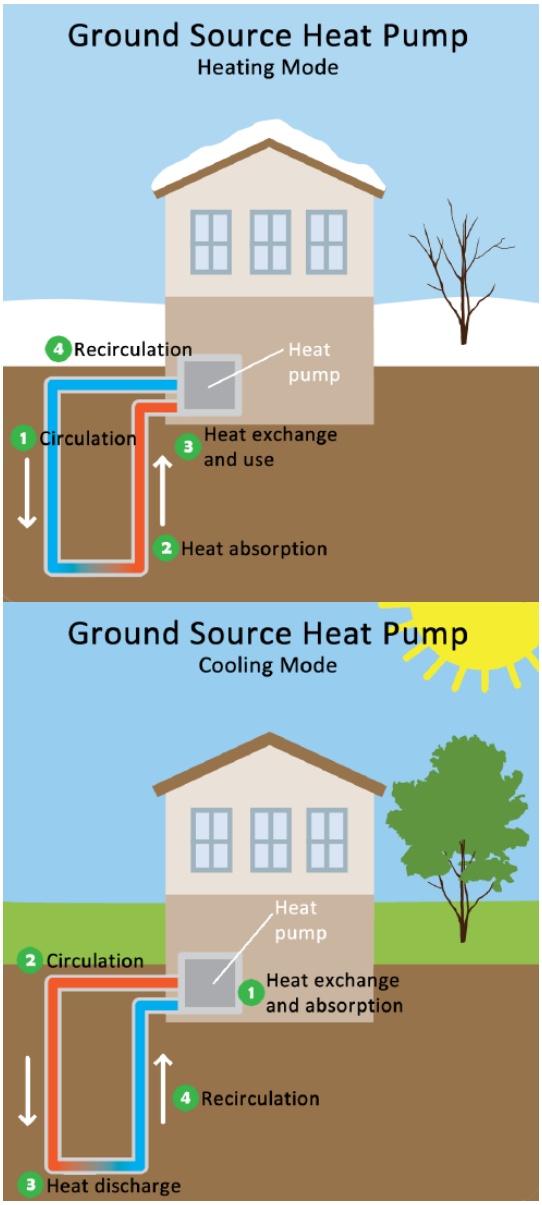
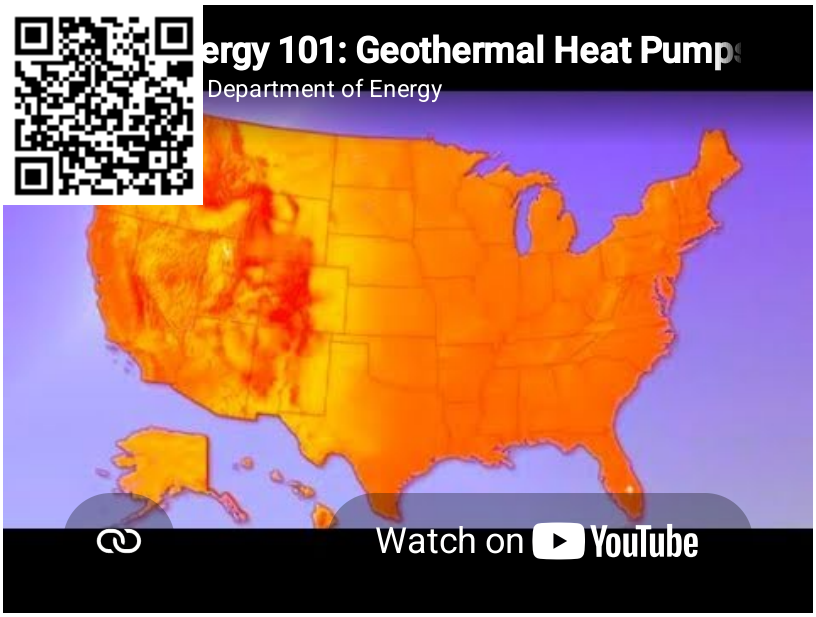


Figure 6.3.4.*d*: A ground source heat pump (geothermal heat pump) in heating mode (during the winter; left) and in cooling mode (during the summer; right). There are four steps for heating mode (left). (1) Circulation: The above-ground heat pump moves water or another fluid through a series of buried pipes or ground loops. (2) Heat absorption: As the fluid passes through the ground loop, it absorbs heat from the warmer soil, rock, or ground water around it. (3) Heat exchange and use: The heated fluid returns to the building where it used for space or water heating. The system uses a heat exchanger to transfer heat into the building's existing air handling, distribution, and ventilation system. (4) Recirculation: Once the fluid transfers its heat to the building, it returns at a lower temperature to the ground loop to be heated again. This process is repeated, moving heat from one point to another. The four steps for cooling mode (right) are similar: (1) heat exchange and absorption, (2) circulation, (3) heat discharge, and (4) recirculation. Images and caption (modified) from EPA (public domain).

This video explains the construction and mechanism of geothermal heat pumps.



Not only does geothermal energy have multiple applications (generating electricity, heating, and cooling), but it is reliable. While solar and wind energy are intermittent, heat is consistently radiated from deep underground. Additionally, the cool temperatures closer to the surface needed for geothermal heat pumps are present year round and at all locations. Geothermal power plants for electricity generation, however, can only be built in specific locations where hot magma is close enough to the Earth's surface. These locations are typically associated with geysers, hot springs, or volcanoes (figure 6.3.4.*e*). Additionally, geothermal power plants are costly to build.



Figure 6.3.4.*e*: Locations of the world's volcanoes. They are concentrated along the west coasts of North American and South America as well as just east of Asia and Australia. These regions, which wrap around the Pacific Ocean, form what is known as the Ring of Fire. Image by Salazar, S.S., Muñoz, Y. & Ospino, A. Analysis of geothermal energy as an alternative source for electricity in Colombia. *Geotherm Energy* **5,** 27 (2017). (CC-BY)

The environmental impact of geothermal energy depends on how it is being used. The use of geothermal heat pumps has almost no negative impact on the environment. Geothermal power plants do not burn fuel to generate electricity, so they generate minimal air pollution. They release less than 1% of the carbon dioxide emissions of a fossil fuel plant. Geothermal plants plants use scrubber systems to clean the air of hydrogen sulfide that is naturally found in the steam and hot water. They emit 97% less sulfur compounds (one cause of acid deposition/acid rain) than are emitted by fossil fuel plants. After the steam and water from a geothermal reservoir have been used, they are injected back into the Earth. One environmental concern associated with geothermal power plants is geothermal drilling during their construction has caused earthquakes, similar to the effects of injection wells for fracking.

Attribution

Modified by Melissa Ha from Renewable Energy and Challenges and Impacts of Energy Use from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY

This page titled 6.3.4: Geothermal Energy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **18.4: Geothermal Energy** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

6.3.5: Hydropower

Hydropower (hydroelectric energy) refers to the energy in moving water, which can be used to generate electricity. It is ultimately driven by solar energy. Water at high elevations is replenished through evaporation (induced by solar heat), condensation, and precipitation. The moving water turns a turbine (figure 6.3.5.*a*) and this moves magnets within a wire coil, causing the movement of electrons in a generator. The result is an electric current (electricity). Combustion of fossil fuels or biofuels, nuclear energy, wind energy, and thermal solar systems all follow this same general mechanism; the difference is what causes the turbine to spin.

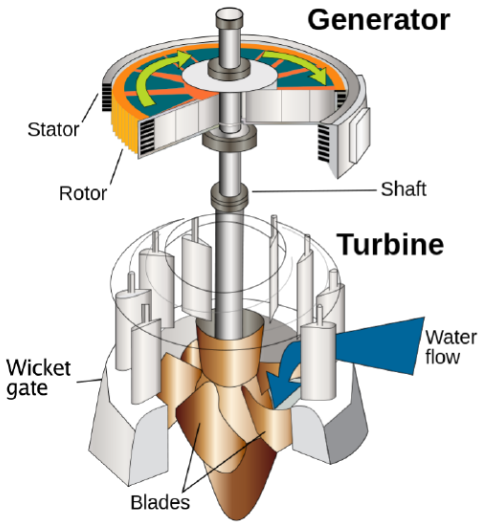


Figure 6.3.5.*a*: A water turbine. Moving water turns the blades of the turbine. The wicket gate controls the water flow. The moving water causes magnets in the rotor to spin and move electrons within the loops of wire in the stator, generating an electric current. Image by U.S. Army Corps of Engineers (public domain).

There are several types of hydropower, but each generates electricity through the same general mechanism described above. The most well-known use of hydropower involves dams. **Dams** block the flow of rivers, producing an artificial lake (reservoir) upstream. The release of water is then controlled. Some of this water moves through a channel (penstock), and this kinetic energy (energy of motion) is harnessed into electricity (figure 6.3.5.*b*). In **run-of-the-river hydroelectricity**, the turbine is placed directly into a river, and the natural flow of water through the river spins the turbine. **Tidal energy** and **energy from ocean waves** (figure 6.3.5.*c*) are also considered forms of hydropower. However, tidal energy arises from the gravitational attraction of the moon and the more distant sun on the earth's oceans, combined with rotation of the earth. In other words, most forms of hydropower are indirect forms of solar energy, with the exception of tidal energy.

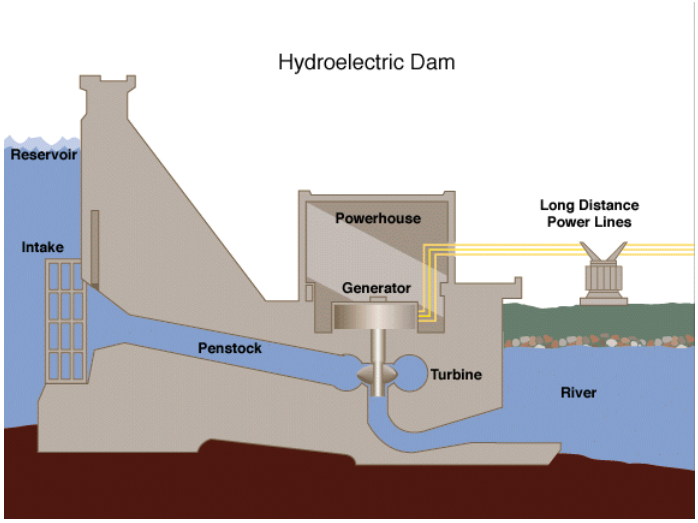


Figure 6.3.5.*b*: A section of a hydroelectric dam. Water in the reservoir passes through the intake into the penstock, which is a narrow channel. The moving water turns a turbine. This powers the generator in the powerhouse. The water continues down the river. Meanwhile electricity is distributed through long distance power lines. Image by Tennessee Valley Authority (public domain).

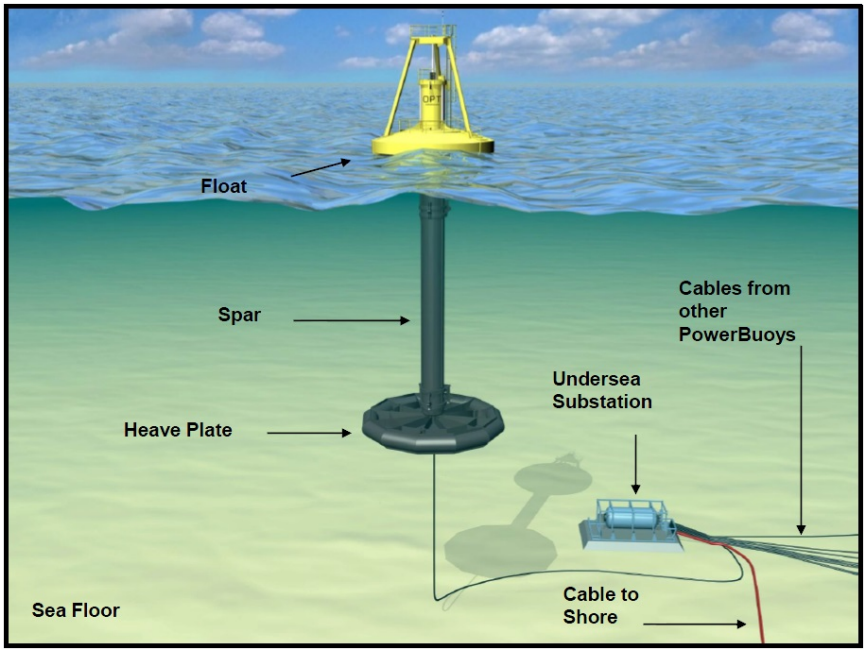


Figure 6.3.5.*c*: One way to harness wave energy is with an absorber. Abosorbers extract energy from the rise and fall of the waves with a buoy. The buoy consists of the float, which is above the water and a heavy plate. The narrow spar connects the two. A cable connects the heavy plate of each buoy to an undersea substation, and another cable runs from the substation to the shore. Image and caption (modified) by OpenEI (public domain).

Small hydropower projects offer power solutions for many remote communities throughout the world, such as those in Nepal, India, China, and Peru, as well as for highly industrialized countries like the United States. **Small hydropower systems** are those that have a capacity between 0.01 to 30 megawatt (MW) of power. For reference, a 1-MW generator operating for an hour (producing 1 MWh) generates enough electricity to powers the average home in the U.S. for a little over a month. Small hydropower systems that generate less than 0.1 MW are more specifically called **microhydropower** systems (figure 6.3.5.*d*). Most of the systems used by home and small business owners would qualify as microhydropower systems. In fact, a 10 kW system generally can provide enough power for a large home, a small resort, or a hobby farm.

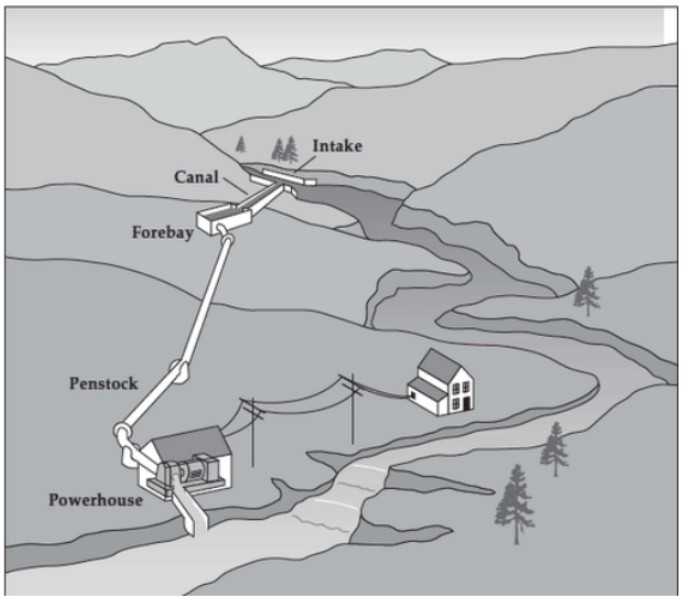


Figure 6.3.5.*d*: A microhydropower system. Although there are several ways to harness the moving water to produce energy, run-of-the-river systems, which do not require large storage reservoirs, are often used for microhydropower projects. For run-of-the-river hydro projects, a portion of a river's water is diverted to a channel, pipeline, or pressurized pipeline (penstock) that delivers it to a waterwheel or turbine. In this image, water from the river moves through the intake then to the canal, forebay, penstock, and powerhouse, which contains the turbine. The moving water rotates the wheel or turbine, which spins a shaft. The motion of the shaft can be used for mechanical processes, such as pumping water, or it can be used to produce electricity. In this image, electrical lines connect the powerhouse to a nearby house.

One benefit of using hydropower is that dams and reservoirs can additionally be used for recreation, flood control, and storing freshwater (see Water Scarcity and Solutions). While we will never run out of moving water, the potential to use hydropower fluctuates based on precipitation. For example, during a drought, water levels in a reservoir decreases, and there is less water released to generate electricity. As with most sources of renewable energy, hydropower is not practical everywhere, and it is most effective in mountainous regions with high rainfall and snowmelt.

Hydropower dams and the reservoirs they create degrades habitat and negatively impacts native species. For example, migration of fish to their upstream spawning areas can be obstructed by dams. In areas where salmon must travel upstream to spawn, such as along the Columbia River in Washington and Oregon, the dams block their way (figure 6.3.5.*e*). This problem can be partially alleviated by using "fish ladders" that help salmon get around the dams (figure 6.3.5.*f*). Fish traveling downstream, however, can get killed or injured as water moves through turbines in the dam. Reservoirs and operation of dams can also affect aquatic habitats due to changes in water temperatures, water depth, chemistry, flow characteristics, and sediment loads, all of which can lead to significant changes in the ecology and physical characteristics of the river both upstream and downstream. When reservoirs first fill with water, they inundate terrestrial (land) habitats, farms, cities, and archeological and cultural sites, sometimes forcing populations to relocate. The terrestrial vegetation slowly decomposes in oxygen-poor conditions, releasing methane into the atmosphere, a potent greenhouse gas. In this sense, hydropower generates few air pollutants during operation, but construction does contribute to climate change.



Figure 6.3.5.*e*: The Grand Coulee Dam on the Columbia River in Washington state created Lake Roosevelt, which provides outdoor recreation such as fishing and swimming. However, it compromised salmon populations and flooded sacred sites for people of the Spokane Tribe. Image by U.S. Bureau of Reclamation (public domain).



Figure 6.3.5.*f*: Left: The fish ladder at John Day Lock and Dam along the Columbia River in Oregon allows adult fish to migrate upstream of the dam on their own. Right: Diagram of a fish ladder, which allows the passage of salmon over a dam. Salmon swim upstream to spawn (reproduce) and are able to swim up the steps of the ladder. The flow from the ladder attracts migrating salmon. At the turning pool, they rotate and then proceed up the next flight of the ladder until they are over the dam. Left image and caption (modified) from U.S. Army Corps of Engineers, and right image by NOAA Fisheries (public domain).

While large-scale dam hydropower projects are often criticized for their impacts on wildlife habitat, fish migration, and water flow and quality, small run-of- the-river projects are free from many of these environmental problems. Because they use the natural flow

of the river, they hardly alter the stream channel and flow. Thus, effects such as oxygen depletion, increased temperature, decreased flow, and impeded upstream migration are not problems for many run-of-the-river projects.

Attribution

Modified by Melissa Ha from Renewable Energy and Challenges and Impacts of Energy Use from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 6.3.5: Hydropower is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **18.5: Hydropower** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

6.3.6: Biofuels (Biomass Energy)

Biofuels (biomass energy) contain energy produced from organisms, such as animal waste, plants, or algae. It is another indirect form of solar energy. Biofuels have many uses. They be burned directly or first converted to ethanol (often with the help of bacteria and fungi) to generate electricity. The heat from combustion produces steam and turns a turbine to power a generator. Biodiesel offers an alternative to petrochemicals for fueling vehicles. Biofuels have even been used to power small planes (figure 6.3.6.*a*). Furthermore, burning wood or straw provides heating.



Figure 6.3.6.*a*: Test flight of the "Green Hornet", a Navy fighter jet partially powered by biofuels. Image by U.S. Navy (public domain).

Unlike fossil fuels, biofuels are **carbon neutral** (figure 6.3.6.*b*). Fossil fuels store carbon that was captured by organisms millions of years ago. When we burn them, carbon dioxide is released much more rapidly than it was removed. Biofuels removed carbon dioxide from the atmosphere more recently, and they form over shorter time spans. When biofuels are burned, this carbon dioxide that was recently removed is released back into the atmosphere.

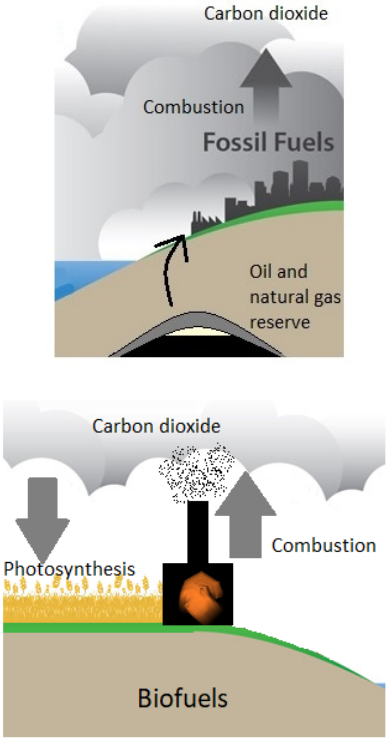


Figure 6.3.6.*b*: Fossil fuels, such as oil and natural gas, contain carbon that has been stored underground for millions of years. Combustion (burning) of these fossil fuels releases carbon dioxide into the atmosphere, contributing to climate change. When biofuels are grown, they capture carbon dioxide through photosynthesis. Burning them or their products for heat, transportation, or electricity generation release this carbon dioxide, which was only recently captured. For this reason, biofuels are considered carbon neutral. Images created by Melissa Ha from NETL/DOE (public domain) and publicdomainvectors.org (public domain).

Another advantage of biofuels is that they can be produced locally and can be cultivated in many different locations. On the other hand, they occupy space that could otherwise be used for food production. To further complicate matters, the characteristics that make a plant species ideal for biofuel (such as being resistant to pests and fast-growing) are also characteristics that help invasive species thrive. Care must taken to contain these species if grown outside of their native ranges.

Combustion of solid municipal waste (see below) or animal wastes as biofuels, reduces waste and generates electricity simultaneously. Unlike most forms of renewable energy, however, combustion of biofuels pollutes the air. (The carbon dioxide released is not an issue since biofuels are carbon neutral, but other air pollutants are also released.) In fact, indoor air pollution from fires used for cooking inside the home is a major cause of death in developing countries.

Each type of biomass must be evaluated for its environmental and social impact in order to determine if it is really advancing sustainability and reducing environmental impacts. For example, cutting down large swaths of forests just for energy production is not a sustainable option because our energy demands are so great that we would quickly deforest the world, destroying critical habitat. For biomass to be a sustainable option, it usually needs to come from waste material, such as lumber mill sawdust, paper mill sludge, yard waste, or oat hulls from an oatmeal processing plant, livestock manure, or trash. These materials would otherwise just accumulate or decompose. Several examples of biofuel use are discussed below in more detail, including the specific advantages and disadvantages of type of use.

Burning Wood

Using wood, and charcoal made from wood, for heating and cooking can replace fossil fuels and may result in lower carbon dioxide emissions. If wood is harvested from forests or woodlots that have to be thinned or from urban trees that fall down or needed be cut down anyway, then using it for biomass does not impact those ecosystems. However, wood smoke contains harmful pollutants like carbon monoxide and particulate matter (see Air Pollution).

For home heating, it is most efficient and least polluting when using a modern wood stove or fireplace insert that are designed to release small amounts of particulates. However, in places where wood and charcoal are major cooking and heating fuels such as in developing countries, the wood may be harvested faster than trees can grow resulting in deforestation (figure 6.3.6.*c*). The largest share of biofuel use comes from traditional biomass, mostly fuel wood gathered for household cooking and heating, often without regard for sustainable replacement.



Figure 6.3.6.*c*: Charcoal production in Uganda. (A) Trees cut down for charcoal. (B) Pieces of firewood. (C) Wood logs read for charcoal burning. (D) Charcoal packed and ready for market. Image and caption (modified) from Bamwesigye et al. *Sustainability* **2020**, 12(20), 8337. (CC-BY)

Biomass can be used in small power plants. For instance, Colgate College has had a wood-burning boiler since the mid-1980s (figure 6.3.6.*d*). In one year it processed approximately 20,000 tons of locally and sustainably harvested wood chips, the equivalent of 1.17 million gallons (4.43 million liters) of fuel oil, avoiding 13,757 tons of emissions and saving the university over $1.8 million in heating costs. The University's steam-generating wood-burning facility now satisfies more than 75% of the campus's heat and domestic hot water needs.



Figure 6.3.6.*d*: Wood chips are one example of a biofuel that can be burned to generate electricity. Source: Ulrichulrich

Municipal Solid Waste

Municipal solid waste (MSW) is commonly known as garbage and can create electricity by burning it directly or by burning the methane produced as it decays. Waste-to-energy processes are gaining renewed interest as they can solve two problems at once: disposal of waste and production of energy from a renewable resource. Many of the environmental impacts are similar to those of a coal plant: air pollution, ash generation, etc. Because the fuel source is less standardized than coal and hazardous materials may be present in MSW, incinerators and waste-to-energy power plants need to clean the gases of harmful materials. The Environmental Protection Agency in the U.S. regulates these plants very strictly and requires anti-pollution devices to be installed. Also, while incinerating at high temperature, many of the toxic chemicals may break down into less harmful compounds. The ash from these plants may contain high concentrations of various metals that were present in the original waste. If ash is clean enough it can be "recycled" as an MSW landfill cover or to build roads, cement blocks, and artificial reefs (similar to coral reefs, but built by humans).

Landfill Gas (Biogas)

Landfill gas (**biogas**) is a sort of human-made "biogenic" gas as discussed above (figure 6.3.6.*e*). Methane is formed as a result of biological processes in sewage treatment plants, waste landfills, anaerobic composting, and livestock manure management systems. This gas is captured and burned to produce heat or electricity. The electricity may replace electricity produced by burning fossil fuels, reducing carbon dioxide emissions. The only environmental impacts are from the construction of the plant itself, similar to that of a natural gas plant.

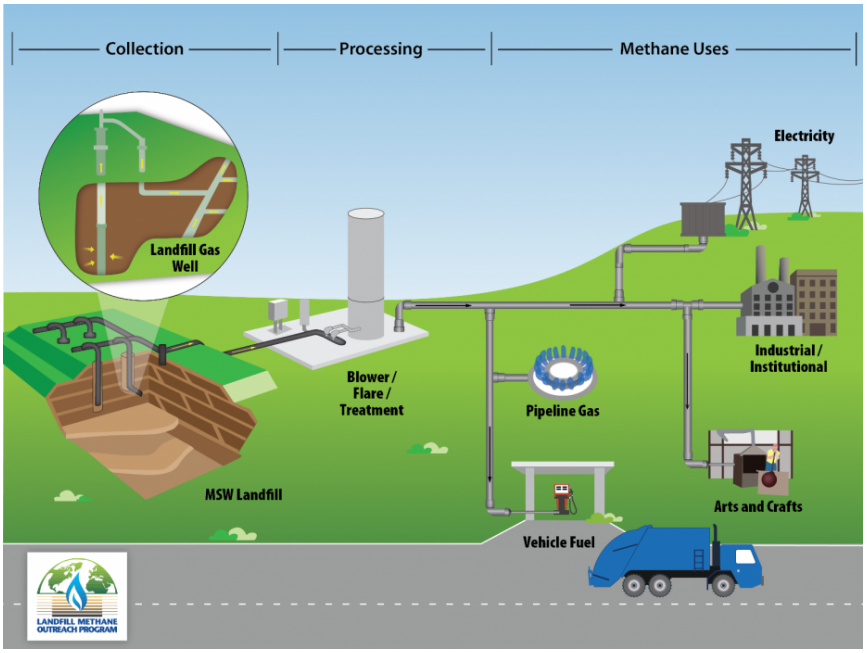


Figure 6.3.6.*e*: The collection and processing of landfill gas to produce methane for multiple uses. First, landfill gas is collected from the landfill gas well. A system of vertical and horizontal piping buried in an municipal solid waste landfill. The landfill gas is then processed and treated for use (blower/flare/treatment). Potential end uses include industrial/institutional uses, arts and crafts, pipeline gas, and vehicle fuel. Image and caption (modified) from Landfill Methane Outreach Program/EPA (public domain).

Bioethanol and Biodiesel

Bioethanol and **biodiesel** are liquid biofuels manufactured from plants, typically crops. Bioethanol can be easily fermented from sugar cane juice, as is done in Brazil. Additionally, it can be fermented from broken down corn starch, as is mainly done in the United States.

The economic and social effects of growing plants for fuels need to be considered, since the land, fertilizers, and energy used to grow biofuel crops could be used to grow food crops instead. The competition of land for fuel versus food can increase the price of food, which has a negative effect on society. It could also decrease the food supply increasing malnutrition and starvation globally. Also, in some parts of the world, large areas of natural vegetation and forests have been cut down to grow sugar cane for bioethanol

and soybeans and palm-oil trees to make biodiesel. This is not sustainable land use. Derived biofuels from parts of plants not used for food, such as stalks, reduces their environmental impact. Biodiesel can be made from used vegetable oil and has been produced on a very local basis. Compared to diesel, a petrochemical derived from crude oil, biodiesel combustion produces less sulfur oxides, particulate matter, carbon monoxide, and unburned and other hydrocarbons, but it produces more nitrogen oxide (see Air Pollution).

Liquid biofuels typically replace petroleum and are used to power vehicles (figure 6.3.6.*f*). Although ethanol-gasoline mixtures burn cleaner than pure gasoline, they also are more volatile and thus have higher "evaporative emissions" from fuel tanks and dispensing equipment. These emissions contribute to the formation of harmful, ground level ozone and smog (see Air Pollution). Gasoline requires extra processing to reduce evaporative emissions before it is blended with ethanol.



Figure 6.3.6.*f*: A bus running on soybean biodiesel. Image and caption (modified) from U.S Department of Energy (public domain).

Attribution

Modified by Melissa Ha from Renewable Energy and Challenges and Impacts of Energy Use from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 6.3.6: Biofuels (Biomass Energy) is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **18.6: Biofuels (Biomass Energy)** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

6.3.7: Energy Conservation

Energy conservation refers to reducing energy waste and increasing efficiency. Energy conservation can involve behavior changes as well as technologies. Some examples of energy conservation have no financial impact. These include turn off and unplugging electronics when not in use, turning down the water heater, and driving efficiently (figure 6.3.7.*a*). Additionally, opening blinds on south-facing windows in the morning during the winter takes advantage of a passive solar technology. Relying on the sun for heating and lighting reduces the use of electricity.



Figure 6.3.7.*a*: A plant on the dashboard of this Ford Fusion (right) indicates if the driver's behavior promotes fuel efficiency. In this picture, there is only one leaf on the plant, indicating low efficiency. Some ways to drive efficiently include following the speed limit and adjusting your speed gradually. Image by ogilvyprworldwide (CC-BY).

Other examples of energy conservation require some financial investment, but they quickly pay for themselves with savings on an energy bill. An **energy audit** is a first step to investigate inefficiencies in one's home. This helps homeowners identify where their home is losing energy, and which problem areas and fixes they should prioritize to save energy and money. For example, an energy audit may reveal places in the home where hot escaping in the winter or entering in the summer. An energy auditor might recommend installing insulation to better seal the home as well as to insulate the hot water heater and pipes. Investing in most high-efficiency appliances also pays for itself relatively quickly (figure 6.3.7.*b*).



Figure 6.3.7.*b*: Energy Star appliances use energy efficiently. Image by MoneyBlogNewz (CC-BY)

This video provides a walkthrough of an energy audit.



Finally, some strategies for energy conservation require sizeable financial investment. They can eventually pay for themselves over extended periods of time. Once example are double-paned, low emissivity (low e) windows (figure 6.3.7.*c*). The two layers of glass trap air between them, which serves as insulation. Additionally, the glass is coated with very small metal dots that allow light to pass through, but infrared (heat) energy is reflected back. If it is hotter outside, heat is emitted back outside; if warmer inside, heat will be emitted back inside. Energy-efficient air conditioners, geothermal heat pumps, and on-demand (tankless) water heaters (figure 6.3.7.*d*) are also examples of energy-conserving technologies that require sizeable investment.

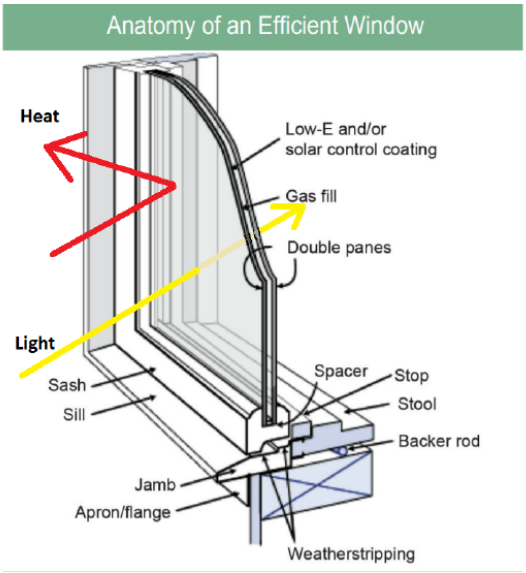


Figure 6.3.7.*c*: Anatomy of an efficient window. A double-paned, low emissivity window allows light from the sun to pass through, but tiny metal dots (not visible) reflect heat (labeled "Low-E and/or solar control coating"). The double panes trap air between them, an insulator (labeled "gas fill"). Spacers are between the glass and the sash, which surrounds the glass. The sash can move out of the frame when the window is open, and weatherstipping prevents air from leaking between the sash and the frame. The stop on the outside prevents the sash swinging outward, and the stool extends perpendicular to the wall outside of the stop. Under it is a backer rod, filling the space between the stool and the wall beneath it. The frame consists of the horizontal sill and vertical jamb. The apron/flange is the trim just below the frame. Image modified from U.S. Department of Energy (public domain).

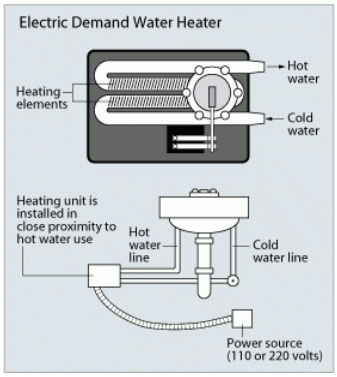


Figure 6.3.7.*d*: An tankless water heater (electric demand water heater). Because hot water is not stored, there is no opportunity for it to cool and need to be reheated. This conserves energy. The heating unit is installed in close proximity to hot water use. It is suppled by a power source (110 or 220 volts). Hot and cold water lines run from the heating unit (which contains heating elements) to the sink. Image by U.S. Department of Energy (public domain).

Attribution

Melissa Ha (CC-BY-NC) and Home Energy Audits. U.S. Department of Energy. Accessed 01-18-2021. (public domain)

This page titled 6.3.7: Energy Conservation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **18.7: Energy Conservation** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

6.3.8: Data Dive- Global Renewable Energy Generation

Overview

Our World in Data (OWID) is a scientific online publication that focuses on using research and data to help tackle global concerns such as poverty, disease, hunger, climate change, war, and inequitable treatment of our world's most vulnerable and unstable communities. Their website is particularly known for publishing a variety of graphs, some even interactive, to present the research that helps explain causes and consequences of global concerns to the public. One example graph, seen below, illustrates renewable energy generation:

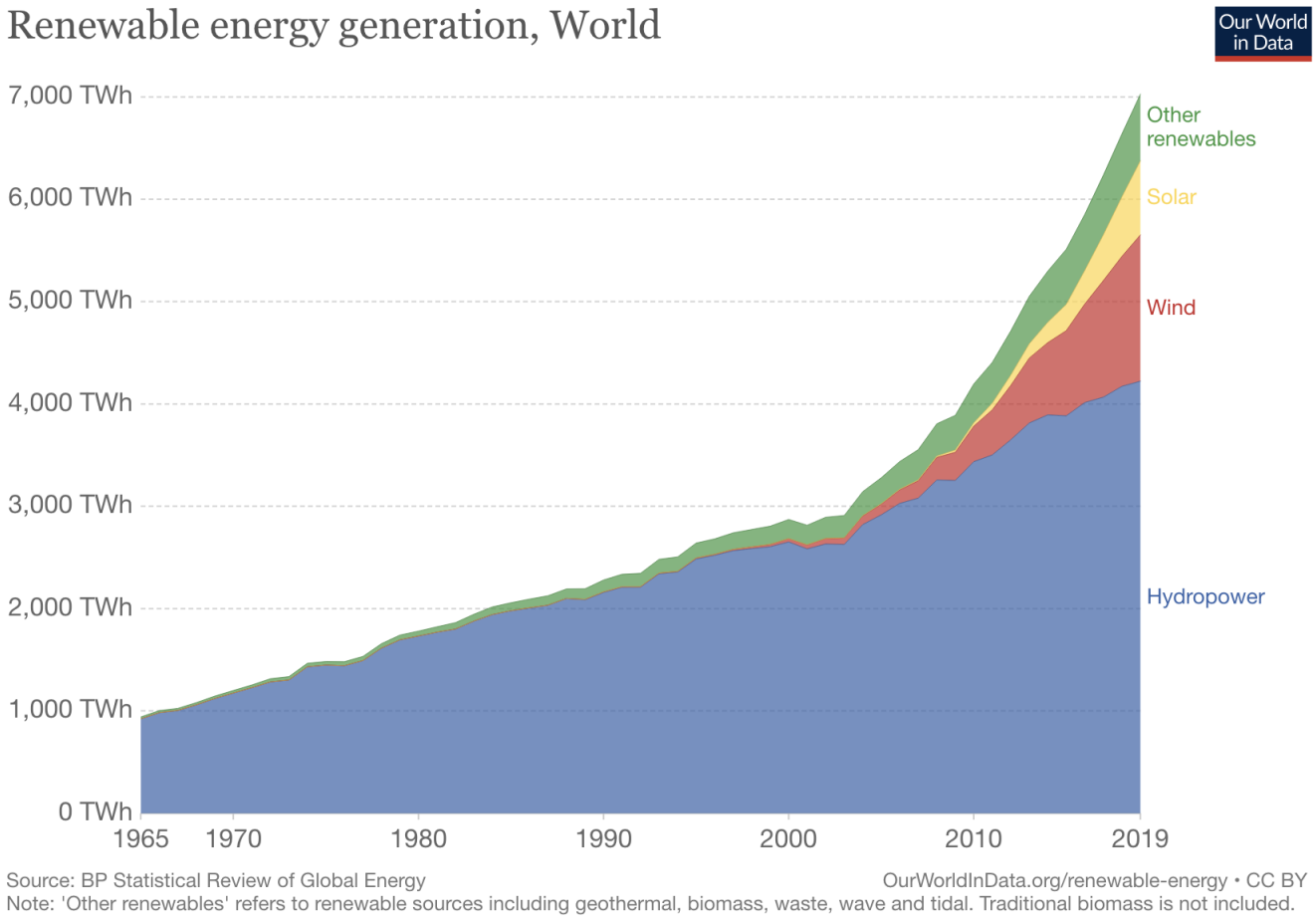


Figure 6.3.8.*a*: Global renewable energy generation. Graph by Our World in Data (CC-BY).

Questions

1. What renewable energy source has had the most consistent use globally?
2. Based on what you have learned in the chapter, what sources do you think are in the "other" category? Don't peek on the website!
3. Describe the pattern you see for the "other" category for 1965 to 2019.
4. Moving into the future, which renewable energy source do you think will grow the most in the next 50 years? Why?

Raw Data For Above Graph(s)

Raw data table link for Our World In Data graph above. (CC-BY)

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 6.3.8: Data Dive- Global Renewable Energy Generation is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **18.8: Data Dive- Global Renewable Energy Generation** by Melissa Ha and Rachel Schleiger has no license indicated.

6.3.9: Review

Summary

After completing this chapter you should be able to...

- Detail the percentage that renewable energy contributes to electricity generation and total energy consumption globally and in the United States.
- Distinguish among the five main types of renewable energy, explain the mechanisms by which they are used, and discuss their advantages and disadvantages.
- Distinguish between passive and active solar technologies and provide examples of each.
- Explain how solar panels generate electricity.
- Distinguish between geothermal power plants and geothermal (ground source) heat pumps.
- Identify the main types of hydropower.
- Explain how environmental damage associated with hydropower can be mitigated.
- Explain why combustion of biofuels is carbon neutral.
- Provide examples of biofuel use and evaluate their sustainability.
- Define energy conservation and provide examples of behaviors and technologies that conserve energy.

Renewable energy sources can be replenished within a short timeframe. This includes **wind**, **solar**, **geothermal**, **hydropower**, and **biofuels**. Each is a direct or indirect form of solar energy with the exception of geothermal energy, which comes from deep underground, and tidal energy, a type of hydropower. Most forms of renewable energy generate electricity by turning a turbine, which powers a generator. For wind energy and hydropower, the motion of air or water, respectively, turns the turbine. For geothermal energy and biofuels, heat from the Earth or from burning organisms, respectively, produce steam, which turns a turbine. Solar panels generate electricity when light energy displaces electrons in semiconductors.

Generally, renewable energy sources generate little air pollution and are not major causes of climate change, but there are exceptions. A downside of renewables is that wind and solar can only be efficiently harnessed in certain locations and are intermittently available. Geothermal energy is more reliable, but geothermal power plants can typically only be built near geysers, volcanoes, or hot springs. Hydropower that involves **dams** and **reservoirs** destroys terrestrial habitats and disrupts aquatic species, but **run-of-the-river hydroelectricity** causes minimal damage. Biofuels are **carbon neutral**, capturing as much carbon dioxide as they release. Biofuels are most sustainable when they are made from materials that would otherwise be wasted.

Energy conservation involves reducing energy use or using it more efficiently. Examples of energy conservation include turning off lights and electronics when they are not in use, driving efficiently, and purchasing energy-efficient applicances.

Attribution

Melissa Ha (CC-BY-NC)

This page titled 6.3.9: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **18.9: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

CHAPTER OVERVIEW

7: Environmental Impacts

The human presence and impact on this Earth has been changing with our population, technology, and affluence. In order to start understanding how these variables have an impact on Earth (in a general sense) the **IPAT** equation was created. As is the case for any equation, IPAT expresses a balance among interacting factors. It can be stated as:

$$I = P \times A \times T$$

where...

- "I" represents the impacts of a given course of action on the environment
- "P" is the relevant human population for the problem at hand
- "A" is the level of consumption per person
- "T" is impact per unit of consumption. Impact per unit of consumption is a general term for technology, interpreted in its broadest sense as any human-created invention, system, or organization that serves to either worsen or uncouple consumption from impact

The equation is not meant to be mathematically rigorous; rather it provides a way of organizing information for a “first-order” analysis. To achieve meaningful reductions of human impact, there are intense debates on where the focus should lie. Where one group sees expensive remedies with little demonstrable return, another sees opportunities for investment in new technologies, businesses, and employment sectors, with collateral improvements in global and national well-being.

This unit will investigate variables that represent the biggest modern components for human impact at local and global scales. Specifically, it will focus on solid waste, pollution, and climate change.

7.1: Solid Waste Management

7.1.1: Waste Generation

7.1.2: Waste Disposal

7.1.3: Solid Waste and Marine Life

7.1.4: Waste Reduction

7.1.5: Data Dive- "Flushable" Wipes

7.1.6: Review

7.2: Pollution

7.2.1: Water Pollution

7.2.1.1: Water Pollutants and Their Sources

7.2.1.2: Water Treatment

7.2.1.3: Mitigating Water Pollution

7.2.2: Air Pollution

7.2.2.1: Ozone Depletion

7.2.2.2: Acid Deposition

7.2.3: Data Dive- Plastic in World Oceans

7.2.4: Review

7.3: Climate Change

7.3.1: The Greenhouse Effect and Climate Change

7.3.2: Data Dive- Flooding in Venice Italy

7.3.3: Review

Attribution

Modified by Rachel Schleiger from Sustainability: A Comprehensive Foundation by Openstax (licensed under CC-BY)

This page titled 7: Environmental Impacts is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

SECTION OVERVIEW

7.1: Solid Waste Management

Chapter Hook

Flushable wipes, are flushable right? No. Over and over, plumbers and waste water treatment workers state that they do not degrade fast enough and clogs drains and sewers across the globe. Not all flushable wipes are created equal, and some are better than others. Unfortunately, there are no legal requirements these products are required to passed to be called flushable, and the only guiding body is the integrity of the company manufacturing the wipes. If you want advice about your body you go to your doctor. If you want advice about your pipes, talk to your plumber. They will tell you that the only things that should go down your toilet are the 3 P's, pee, poo, and (toilet) paper.



Figure 7.1.*a*: Wipe isle at the store. Image by 維基小霸王 in Wikimedia Commons (CC-BY-SA4.0)

7.1.1: Waste Generation

7.1.2: Waste Disposal

7.1.3: Solid Waste and Marine Life

7.1.4: Waste Reduction

7.1.5: Data Dive- "Flushable" Wipes

7.1.6: Review

Attribution

Modified by Rachel Schleiger (CC-BY-NC).

This page titled 7.1: Solid Waste Management is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

7.1.1: Waste Generation

In natural systems, there is no such thing as waste. Everything flows in a natural cycle of use and reuse. Living organisms consume materials and eventually return them to the environment, usually in a different form, for reuse. **Solid waste** (or trash) is a human concept. It refers to a variety of discarded materials, not liquid or gas, that are deemed useless or worthless (figure 7.1.1.*a*). Humans modify natural substances, making them more difficult to breakdown, or store matter under conditions that slow its degradation. However, what is worthless to one person may be of value to someone else, and solid wastes can be considered misplaced resources. Learning effective ways to reduce the amount of wastes produced and to recycle valuable resources contained in the wastes is important if humans wish to maintain a livable and sustainable environment.



Figure 7.1.1.*a*: Trash is a uniquely human concept. Image by Piqsels (public domain).

Types of Waste

There are many subcategories of waste (figure 7.1.1.*b*). **Biodegradable** waste can be broken down by microbes whereas **nondegradable** waste does not readily break down. **Hazardous waste** is trash that presents a health risk. Specifically, hazardous wastes are defined as materials which are toxic, **carcinogenic** (cause cancer), **mutagenic** (cause DNA mutations), **teratogenic** (cause birth defects), highly flammable, corrosive, or explosive. Examples include batteries, fluorescent lights, various cleaners, and **e-waste**, which arises from discarded electronics. Precious metals can be extracted and recycled from hazardous waste, but it must be done safely.



Figure 7.1.1.*b*: Plastic cutlery is nondegradable, but these spoons and forks and made of wood, a biodegradable material. Image by BRRT/Pixabay (Pixabay license).

Solid Waste Sources and Composition

Agriculture and industry generate most of the waste globally followed by mining (figure 7.1.1.*c*). Together, these generate **non-municipal solid waste**. Some common items that are classified as non-municipal waste are: construction materials (roofing shingles, electrical fixtures, bricks); waste-water sludge; incinerator residues; ash; scrubber sludge; oil/gas/mining waste; railroad ties, and pesticide containers. The remaining waste, **municipal solid waste** (MSW), is made up of discarded solid materials from

residences, businesses, and city buildings. Globally, waste generated from industry generated was nearly 18 times that of MSW, and waste generated from agriculture was more than 4.5 times that of MSW.

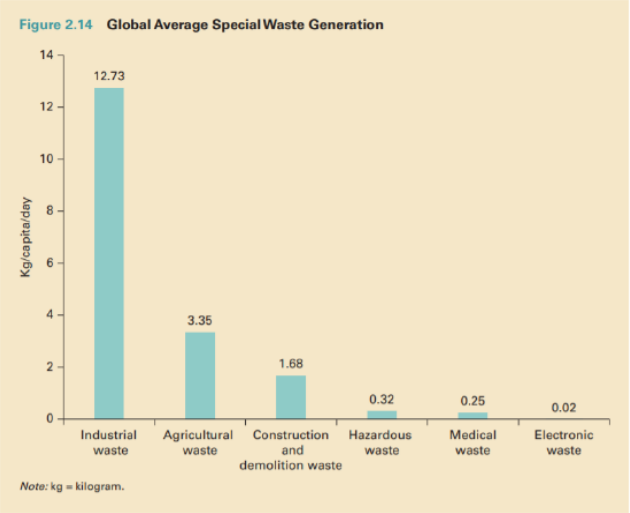


Figure 7.1.1.*c*: Global average special waste generation showing selected categories of waste in kilograms (kg) per person (capita) per day. 1 kg = 2.2 lbs. In order of greatest to least waste generation is industrial waste (12.73 kg/capita/day), agricultural waste (3.35), construction and demolition waste (1.68), hazardous waste (0.32), medical waste (0.25), and electronic waste (0.02). Image from What a Waste 2.0 : A Global Snapshot of Solid Waste Management to 2050 (CC-BY)

The world generates 2.21 billion tons of municipal solid waste annually. Municipal solid waste makes up 3-9% of all waste in the U.S., and 292.4 million tons were generated in U.S. in 2018 (4.9 lbs per person per day). This contrasts with average values in Sub-Saharan Africa (1.01 lbs per person per day), South Asia (1.15 lbs per person per day), and East Asia and the Pacific (1.23 lbs per person per day). Daily waste generation per person in high income countries is projected to increase by 19 percent by 2050, compared to low- and middle-income countries where it is anticipated to increase by approximately 40 percent or more.

Municipal solid waste consists of materials from plastics to food scraps. The most common waste product is paper (about 23% of the total; figure 7.1.1.*d*). Other common components are yard waste (green waste), plastics, metals, wood, glass and food waste. The composition of the municipal wastes can vary from region to region and from season to season.

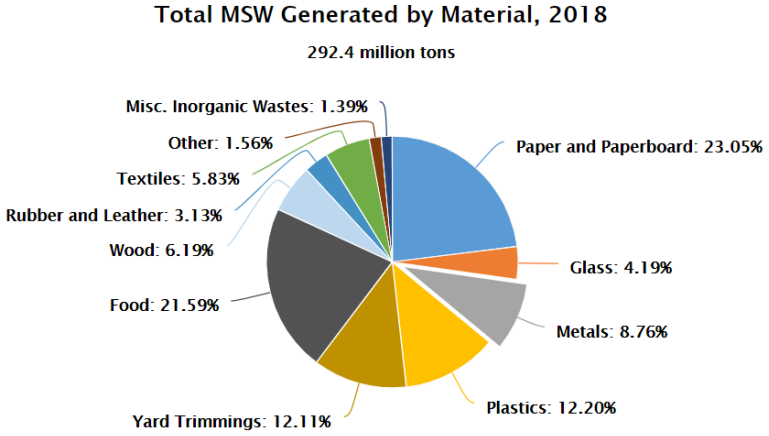


Figure 7.1.1.*d*: Composition of municipal solid waste (MSW) in the U.S. in 2018 before recycling. Paper and paperboard comprise 23.05% followed by food (21.59%), plastics (12.20%), yard trimmings (12.11%), metals (8.76%), wood (6.19%), textiles (5.83%), glass (4.19%), rubber and leather (3.13%), other (1.56%), and miscellaneous inorganic waste (1.39%). Image from EPA (public domain).

Attribution

Modified by Melissa Ha from the following sources:

- Systems of Waste Management from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors (CC-BY). Download for free at CNX.
- Solid Waste form *AP Environmental Science* by University of California College Prep (CC-BY). Download for free at CNX.
- Kaza, Silpa; Yao, Lisa C.; Bhada-Tata, Perinaz; Van Woerden, Frank. 2018. What a Waste 2.0 : A Global Snapshot of Solid Waste Management to 2050. Urban Development;. Washington, DC: World Bank. © World Bank. (CC-BY)

This page titled 7.1.1: Waste Generation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **19.1: Waste Generation** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

7.1.2: Waste Disposal

There are three primary methods for waste disposal: open dumps, sanitary landfills, and incineration. Sanitary landfills and incineration prevent reuse, recycling, and proper decomposition. While open dumps promote decomposition better than other methods of waste disposal and allow discarded materials to be salvaged or recycled, they promote disease spread and cause water pollution. They are thus illegal in many countries.

Open Dumps

Open dumps involve simply piling up trash in a designated area and is thus the easiest method of waste disposal (figure 7.1.2.*a*). Open dumps can support populations of organisms that house and transmit disease (reservoirs and vectors, respectively). Additionally, contaminants from the trash mix with rain water forming **leachate**, which infiltrates into the ground or runs off. This liquid leachate may contain toxic chemicals such as dioxin (a persistent organic pollutant), mercury, and pesticides.



Figure 7.1.2.*a*: An open dump in Vietnam. Image by Julien Belli (CC-BY).

Sanitary Landfills

After recycling, composting, and incineration, the remaining 50% of municipal solid waste (MSW) in the U.S. was discarded in **sanitary landfills** (figure 7.1.2.*b*). Trash is sealed from the top and the bottom to reduce contamination of surroundings (figure 7.1.2.*c*). Rainwater that percolates through a sanitary landfill is collected in the bottom liner, and this bottom layer thus prevents contamination of groundwater. The groundwater near the landfill is closely monitored for signs of contamination from the leachate. Layers of soil on top prevent disease spread. Each day after garbage is dumped in the landfill, it is covered with clay or plastic to prevent redistribution by animals or the wind.



Figure 7.1.2.*b*: A sanitary landfill in 1972. Image by Bill Shrout/EPA (public domain).

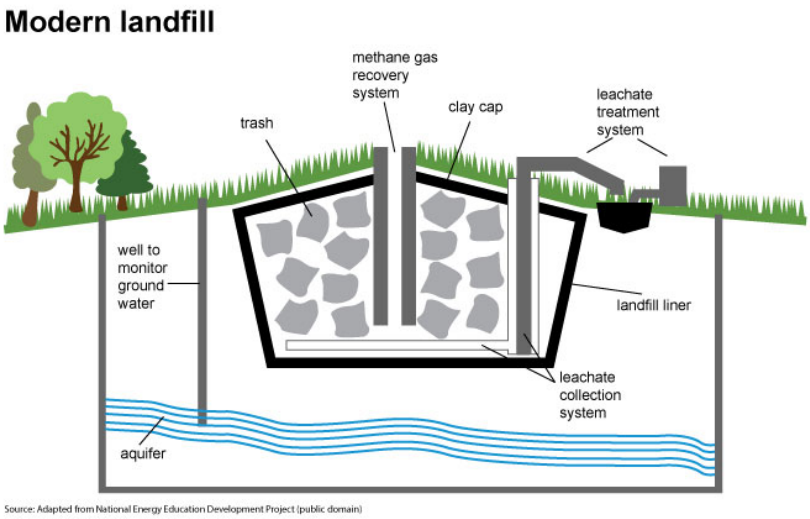


Figure 7.1.2.*c*: A sanitary landfill. Trash is compacted and stored underground. A landfill liner prevents leachate from draining into the aquifer and contaminating groundwater, which is monitored using a well. Additionally, leachate is collected and treated (leachate treatment system). A clay cap prevents animals or wind from accessing the trash. A methane gas recovery system captures the potent greenhouse gas, which can be used as a biofuel. Image by EPA/National Energy Education Development Project (public domain).

Several practices can reduce the environmental impact of sanitary landfills. Compacting in landfills reduces water and oxygen levels, slowing decomposition and promoting methane release. In the U.S., the Clear Air Act requires that landfills of a certain size collect landfill gas (biogas), which can be used as a biofuel for heating or electricity generation. Other gases such as ammonia and hydrogen sulfide may also be released by the landfill, contributing to air pollution. These gases are also monitored and, if necessary, collected for disposal. To address the often dry condition of wastes within landfills, the concept of bioreactor landfills has emerged. These recirculate leachate and/or inject other liquids to increase moisture and promote decomposition (and therefore increasing the rate of biogas production). Upon closure, many landfills undergo "land recycling" and can be redeveloped as golf courses, recreational parks, and other beneficial uses.

With respect to waste mitigation options, landfilling is quickly evolving into a less desirable or feasible option. Landfill capacity in the United States has been declining for several reasons. Older existing landfills are increasingly reaching their authorized capacity. Additionally, stricter environmental regulations have made establishing new landfills increasingly difficult. Finally, public opposition delays or, in many cases, prevents the approval of new landfills or expansion of existing facilities.

Incineration

Incineration is simply burning trash. This has several advantages: it reduces volume and can be used to generate electricity (**waste-to-energy**). In fact, the sheer volume of the waste is reduced by about 85%. Incineration is costly, however, and it pollutes air and water. Air pollutants released by incineration include particulates, sulfur dioxide, nitrogen oxides, methane, heavy metals (such as lead and mercury), and dioxins. The byproduct of incineration, ash, is often toxic. Depending on its composition, ash might require special disposal; other types of ash can be repurposed.

An incinerator processes trash and burns it in a **combustion chamber** (figure 7.1.2.$d-e$). The heat boils water, and the resultant steam is used to generate electricity. The smoke (called flue gases) goes through a **pollution-removal** before it is released, but it still contains some pollutants. The U.S. incinerated 11.8% of MSW in 2018.



Figure 7.1.2.*d*: Wheelabrator Technologies' Waste-to-Energy plant in Saugus, Massachusetts has been in service since 1975. Image and caption (modified) from Fletcher6 (CC-BY).

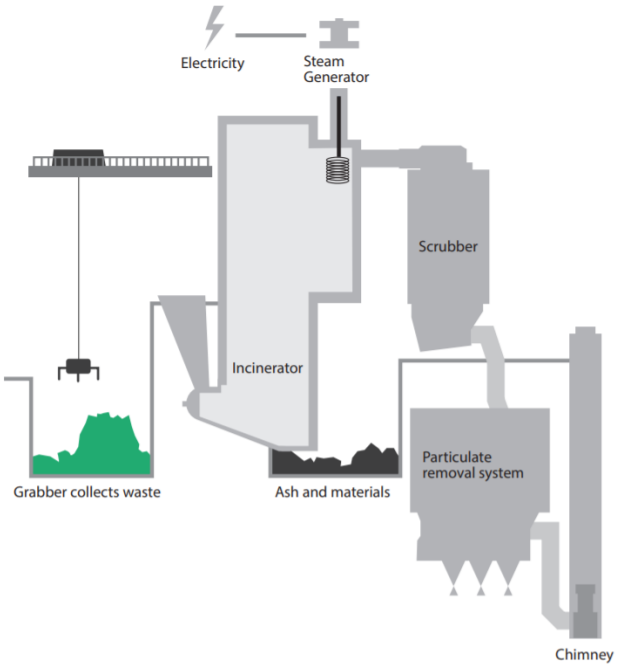


Figure 7.1.2.*e*: A schematic diagram of a waste incineration facility (waste-to energy system). A grabber collects the waste and transfers it into the incinerator. Heat from combusting the trash is used to produce steam and generate electricity. Emissions (flue gases) pass through a scrubber and particulate removal system to limit pollution before the gases are released through a chimney. Ash remains after incineration. Image by Kaza, Silpa; Bhada-Tata, Perinaz. 2018. Decision Maker's Guides for Solid Waste Management Technologies. Urban Development Series Knowledge Papers;. World Bank, Washington, DC. © World Bank. (CC-BY)

There are two kinds of waste-to-energy systems: mass burn incinerators and refuse-derived incinerators. In **mass burn incinerators** all of the solid waste is incinerated. The heat from the incineration process is used to produce steam. This steam is used to drive electric power generators. Acid gases from the burning are removed by chemical scrubbers. Any particulates (small particles that remain suspended in the air) in the combustion gases are removed by **electrostatic precipitators**, which charge particulates and remove them with electrodes. The cleaned gases are then released into the atmosphere through a tall stack. The ashes from the combustion are sent to a landfill for disposal.

It is best if only combustible items (paper, wood products, and plastics) are burned. In a **refuse-derived incinerator**, non-combustible materials are separated from the waste. Items such as glass and metals may be recycled. The combustible wastes are

then formed into fuel pellets which can be burned in standard steam boilers. This system has the advantage of removing potentially harmful materials from waste before it is burned. It also provides for some recycling of materials.

Attribution

Modified by Melissa Ha from the following sources:

- Systems of Waste Management from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors (CC-BY). Download for free at CNX.
- Solid Waste from *AP Environmental Science* by University of California College Prep (CC-BY). Download for free at CNX.
- Biomass explained: Landfill gas and biogas. 2020. EIA. Accessed 01-18-2021 (public domain).

This page titled 7.1.2: Waste Disposal is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **19.2: Waste Disposal** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

7.1.3: Solid Waste and Marine Life

Ocean dumping has also been a popular way for coastal communities to dispose of their solid wastes. In this method, large barges carry waste out to sea and dump it into the ocean (figure 7.1.3.*a*). That practice is now banned in the United States due to the pollution problems it created. However, much trash escapes to the ocean through littering and illegal dumping. Additionally, wind carries trash off of overfilled containers, landfills, or open dumps (where they are still used).



Figure 7.1.3.*a*: Left: A garbage barge leaves Manhattan. Right: Marine debris at the beach. Left image and caption by EPA (public domain), and right image and caption by NOAA (public domain).

Garbage patches form from waste that escapes to the ocean. Plastic breaks down into smaller pieces when exposed to sunlight and ample oxygen. These pieces of plastic are trapped in calm parts of the ocean that are surrounded by strong, circular currents (figure 7.1.3.*b*). Garbage patches may not be visible from the surface of the ocean, and the plastic pieces may be up to 20 meters deep.

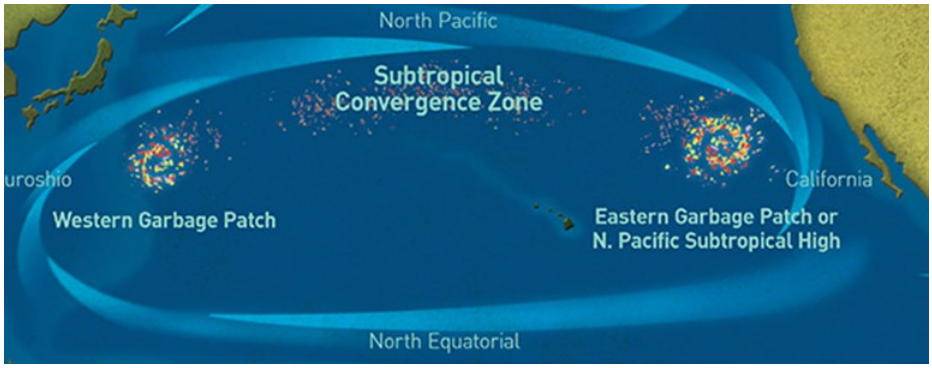


Figure 7.1.3.*b*: Locations of the largest garbage patches in the Pacific Ocean. The Western Garbage Patch is closer to Japan, and the Eastern Garbage Patch is closer to California. The Subtropical Convergence Zone is in the middle to the north. Ocean currents including the Kuroshio Current (left), North Pacific Current (top), California Current (right), and North Equatorial Current (bottom) are labeled. Image by Boyan Slat (public domain).

Plastic and other trash harms marine life in several ways. Old fishing nets and plastic six-pack rings trap wildlife. Additionally, organisms that ingest plastic risk choking. Ingested plastic may cut internal organs, limit space available for food, and carry toxins (figure 7.1.3.*c*).



Figure 7.1.3.*c*: Much of this dead albatross's digestive system was filled with plastic waste. Image by Forest and Kim Starr (CC-BY).

Attribution

Melissa Ha (CC-BY-NC) and Solid Waste from *AP Environmental Science* by University of California College Prep (CC-BY). Download for free at CNX.

This page titled 7.1.3: Solid Waste and Marine Life is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **19.3: Solid Waste and Marine Life** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

7.1.4: Waste Reduction

The **waste management hierarchy** lists processes to address waste in order of preference with the goals of limiting waste, minimizing environmental impact, and maximizing additional benefits (such as electricity generation or production of recycled goods). It is symbolized as an upside triangle, meaning that the largest volume of waste should be handled using the first process (most preferred), and the smallest volume should be handled using the last process (least preferred). These processes in order of preference are reduction, recycling, energy recovery (such as incineration), treatment, and disposal (figure 7.1.4.*a*). Individuals can exercise the first two processes by applying the four R's (see below). Unfortunately, the U.S. and the world overall, handles a large volume of waste using the least preferred method (disposal). As previously mentioned, about half of municipal solid waste (MSW) goes to a sanitary landfill for disposal. In contrast, only 23.5% of municipal solid waste (MSW) was recovered and recycled and another 8.5% was composted in 2018. Waste stream percentages also vary widely by region. As an example, San Francisco, California recycles or composts 80% of its waste material, whereas Houston, Texas recycles or composts only 19%.

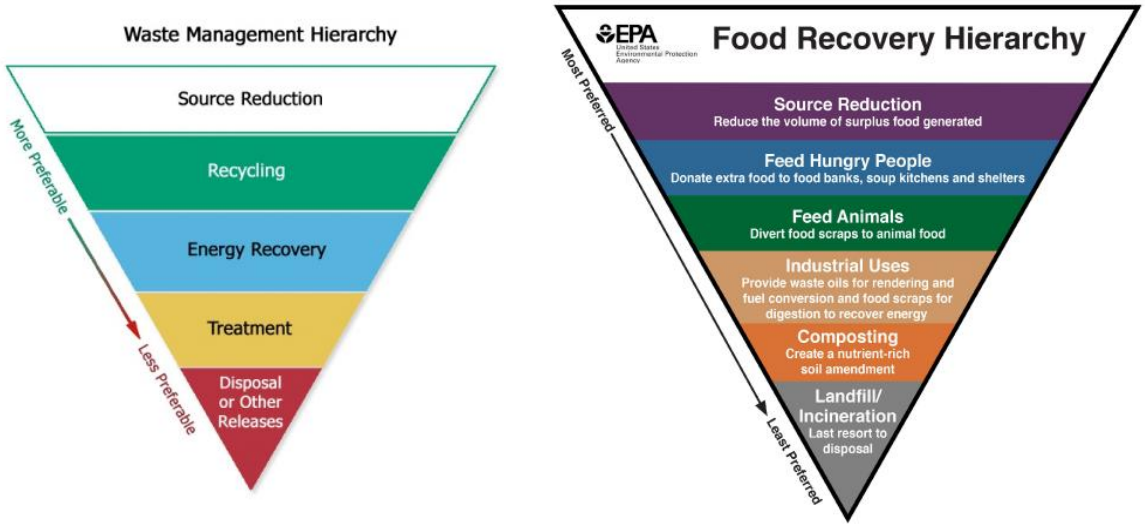


Figure 7.1.4.*a***:** Left: The waste management hierarchy lists processes for handling waste from most to least preferable. They are source reduction, recycling, energy recovery, treatment, and disposal or other releases. Right: The food recovery hierarchy is a waste management hierarchy specific to food waste. It begins with source reduction (reduce the volume of surplus food generated). Next is feed hungry people (donate extra food to food banks, soup kitchens, and shelters. Third is feed animals (divert food scraps to animal food. Fourth is industrial uses (provide waste oils for rendering and fuel conversion and food scraps for digestion to recover energy. Fifth is composting (create a nutrient-rich soil amendment). The remaining waste goes to landfill or incineration as a last resort to disposal. Images by EPA (public domain).

Source reduction (waste minimization) refers to the strategies that minimize the amount of generated waste and/or reduce the toxicity of the resultant waste when designing or manufacturing products or services. Waste can be reduced by reusing materials, using less hazardous substitute materials, or by modifying components of design and processing. Source reduction in manufacturing not only saves resources, but it also reduces costs for the manufacturer and indirectly for the consumer. For example, minimal packaging reduces material use, increases distribution efficiency, and reduces fuel consumption and resulting air emissions. Similarly, building materials can be designed to reduce the overall mass of material needed for a given structure.

Recycling refers to recovery and reprocessing of useful materials. Numerous examples of successful recycling and reuse efforts are encountered every day. Many construction materials can be reused, including concrete, asphalt materials, masonry, and reinforcing steel. "Green" plant-based wastes are often recovered and immediately reused for mulch or fertilizer applications. Many industries also recover various byproducts for reuse. In some cases, the recycled materials are used as input materials and are heavily processed into end products. Common examples include the use of scrap paper for new paper manufacturing, or the processing of old aluminum cans into new aluminum products. In other cases, reclaimed materials undergo little or no processing prior to their

reuse. Some common examples include the use of tree waste as wood chips, or the use of brick and other fixtures into new structural construction. In any case, the success of recycling depends on effective collection and processing of recyclables, markets for reuse, and public acceptance and promotion of recycled products and applications utilizing recycled materials.

The Four R's

The **four R's** (refuse, reduce, reuse, and recycle) are strategies that allow individuals to limit the volume and negative impacts of the waste they generate. They resemble the waste management hierarchy, but focus at the individual scale rather than the scale of a company or waste management system. They are listed in order of most to least environmentally beneficial, and they are all better alternatives than discarding trash into the landfill. Recycling is last because it requires energy to reprocess waste.

To **refuse** is to decline products or processes that harm the environment if you do not need them. Choosing products without packaging is an example of refusing. To **reduce** is to choose products or processes with lower **ecological footprint** (the area of land required to produce them). Examples include products with limited packaging or durable items rather than flimsy or disposable ones. **Reuse** refers to using a product multiple times or finding an alternative use for it. For example, one could share, borrow, or rent items. A plastic yogurt or pudding container could be repurposed for storage or gardening. As noted above, **recycling** means to return used items to be reprocessed (figure 7.1.4.*b*). Commonly recycled materials include cardboard, glass, cans, and certain plastics.



Figure 7.1.4.*b***:** Recycling, or returning materials to be reprocessed, is the last of the four R's. If items cannot be refused, reduced, or reused, recycling is a better alternative than discarding into the landfill. Image by Intel Free Press (CC-BY).

Resin identification codes are the triangular symbols on recyclable plastics that indicate their composition (figure 7.1.4.*c*). A variety of materials make up plastics, and they are not all recycled the same way. Furthermore, your local recycling service may recycle some types of plastics but not others. Plastics with codes 1 and 2 are commonly recycled. For the codes, check with your local recycling service. Local businesses may house collection bins for plastics that are not commonly recycled. For example, grocery stores commonly have receptacles for recycling plastic bags, which fall under code 4.

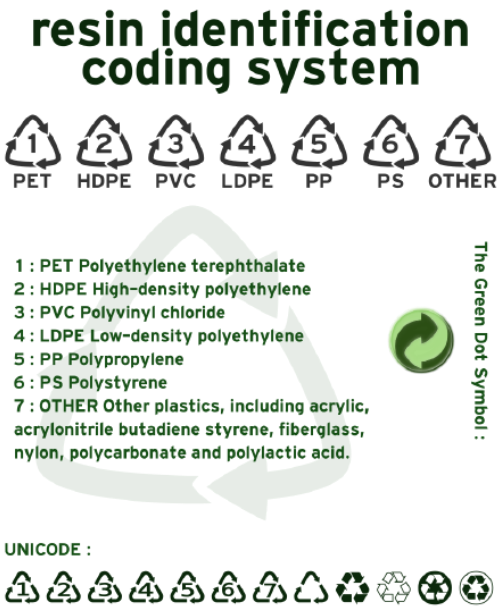


Figure 7.1.4.*c*): The resin identification coding system explains the chemical composition of various plastics. Some are commonly recycled, and others are not. The text in the image is, "1: PET Polyethylene terephthalate, 2: HDPE High-density polyethylene, 3: PVC Polyvinyl chloride, 4: LDPE Low-density polyethylene, 5: PP Polypropylene, 6: PS Polystyrene, 7: OTHER Other plastics, including acrylic, acrylonitrile butadiene styrene, fiberglass, nylon, polycarbonate and polylactic acid. Image by Filtre/OpenClipart Library (public domain).

Composting

Compost is organic material that can be added to soil to help plants grow (figure 7.1.4.*d*). Food scraps and yard waste together currently make up more than 30% of what we throw away and could be composted instead. Making compost keeps these materials out of landfills where they take up space and release methane, a potent greenhouse gas that contributes to climate change. Furthermore, compost enriches soil, helping retain moisture and suppress plant diseases and pests. It reduces the need for chemical fertilizers and encourages the production of beneficial bacteria and fungi. Composting can be done at an industrial scale, such as with yard waste collected from homes in a city. It can also be done at home.



Figure 7.1.4.*d*): Compost is a nutrient-rich substance formed from decomposition of yard and food waste that enriches garden soil. Image by EPA (public domain).

All composting requires three basic ingredients: browns, greens, and water. "Browns" are materials such as dead leaves, branches, and twigs. "Greens" are materials such as grass clippings, vegetable waste, fruit scraps, and coffee grounds. The brown materials provide carbon for your compost, the green materials provide nitrogen, and the water provides moisture to help break down the organic matter. A compost pile should have an equal amount of browns to greens. Layers of these organic materials with different particle sizes should be alternated.

Backyard Composting

Backyard composting requires a compost bin or pile to be placed in dry, shady spot near a water source (figure 7.1.4.*e*). Brown and green materials are added as they are collected, and larger pieces are chopped or shredded before they are added to the pile. Dry materials are moistened as they are added. Once the compost pile is established, grass clippings and green waste are mixed into the pile when added. When fruit and vegetable wastes are added, they are buried under 10 inches of compost material. Some homeowners cover the top of the compost with a tarp to keep it moist. When the material at the bottom is dark and rich in color, the compost is ready to use. This usually takes anywhere between two months to two years.



Figure 7.1.4.*e*): This compost bin contains food scraps and yard waste. Image by Ben_Kerckx/Pixabay (Pixabay license)

Indoor Composting

For those who do not have the space for an outdoor compost pile, materials can be composted indoors using a special type of bin, which can be purchased at a local hardware store, gardening supplies store, or made at home. Care must be taken to properly manage the compost pile and keep track of what is added such that it does not attract pests or rodents or smell bad. The compost should be ready in two to five weeks.

Attribution

Modified by Melissa Ha from the following sources:

- Systems of Waste Management from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors (CC-BY). Download for free at CNX.
- Composting at Home. Environmental Protection Agency. Accessed 01-18-2021 (public domain).

This page titled 7.1.4: Waste Reduction is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **19.4: Waste Reduction** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

7.1.5: Data Dive- "Flushable" Wipes

Overview

Ryerson University published a 2019 report titled, Defining “Flushability” for Sewer Use. This publication had the objective of making the public think twice about what you flush down the toilet in order to save money and the environment. In this experiment they tested 101 single use products, some labeled (and some not labeled) flushable. Some of the results of this can be seen in the graph below:

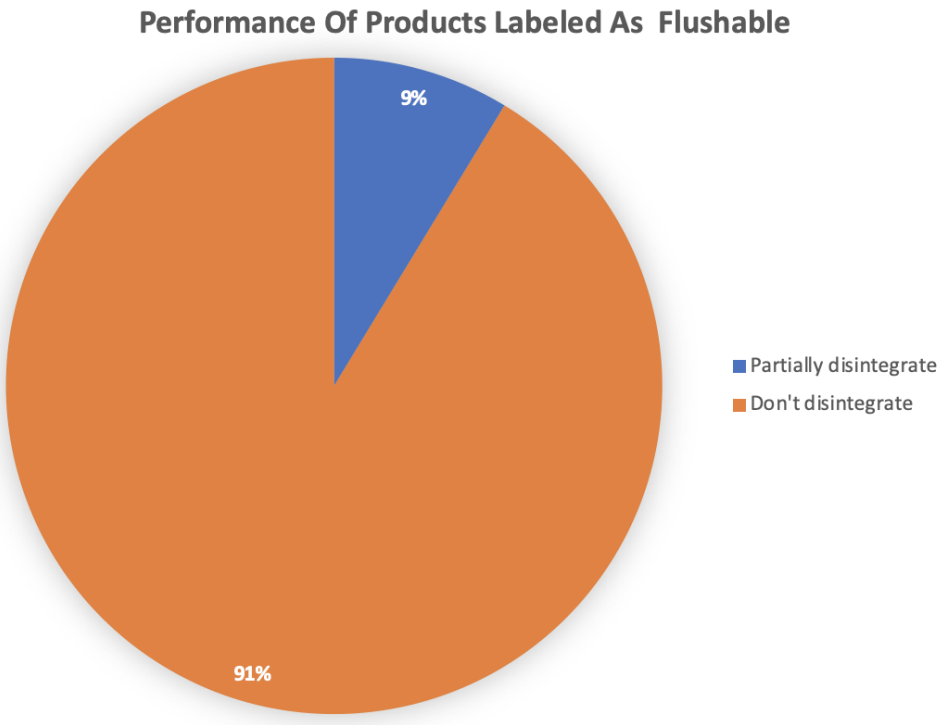


Figure \(\PageIndex{a}\): Percent of flushable wipes that could pass as flushable. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Khan A, Orr B, Jocksimovie D 2019 Flushability Report.

Questions

1. What type of data is this (qualitative or quantitative)?
2. What question(s) are the authors trying to answer with this graph?
3. What result is this graph illustrating relative to products labeled as “flushable?”
4. Provide a brief reflection for how these results make you feel and what they make you think about?
5. Do you think there should be a policy put in place that places restrictions on what the word “flushable” means? Why?

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 7.1.5: Data Dive- "Flushable" Wipes is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **19.5: Data Dive- "Flushable" Wipes** by Melissa Ha and Rachel Schleiger has no license indicated.

7.1.6: Review

Summary

After completing this chapter you should be able to...

- Distinguish between biodegradable and nondegradable waste.
- Provide examples of hazardous waste.
- List the major waste generating processes globally.
- Compare the volume of waste generated per person per day in the U.S. with other countries around the world.
- Define municipal solid waste and describe its composition in the U.S.
- Compare open dumps, sanitary landfills, and incineration, outlining the advantages and disadvantages of each approach.
- Explain how trash enters the ocean and how it harms marine life.
- Explain the waste management hierarchy.
- List the four R's and provide examples of each.
- Explain the benefits and process of composting.

Solid waste refers to a variety of discarded materials that are deemed useless and discarded by humans. Waste may be easily broken down (**biodegradable**) or **nondegradable**. **Hazardous waste** is harmful to human health. Examples include batteries, cleaners, and **e-waste**. Agriculture, industry, and mining generate most of the world's waste. **Municipal solid waste** (MSW) is discarded from residences, businesses, and city buildings, and the U.S. generates more MSW per person per day than most countries around the world.

Three strategies for disposing of waste are open dumps, sanitary landfills, and incinerators. **Open dumps** cause disease spread, air pollution, and water pollution, and are thus illegal in the U.S. About half of the MSW in the U.S. goes to **sanitary landfills**, which compact trash and seal it to prevent pollution. **Incinerators** burn trash, which reduces its volume while generating energy.

Ocean dumping and escape of trash from overfilled bins and open dumps has contaminated the ocean with many plastics. These break into small pieces and form **garbage patches**. Plastic in the ocean can trap, poison, or otherwise harm marine life.

The **waste management hierarchy** diagrams the priorities of waste reduction. Source reduction and recycling are key to limiting waste and mitigating its environmental impacts. However, a large volume of trash goes to disposal, the least preferred method for handling waste. Individuals can limit their own waste by employing the **four R's**: **refuse**, **reduce**, **reuse**, and **recycle**. **Composting** at home also reduces food and yard waste.

Attribution

Melissa Ha (CC-BY-NC)

This page titled 7.1.6: Review is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **19.6: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

7.2: Pollution

Chapter Hook

In the midst of the COVID-19 pandemic many people in the United States were feeling the fatigue of wearing face masks for protection. However, people in the United States should consider themselves lucky that normally their air is clean enough that face masks are only really necessary for people who have sensitive immune systems (to things like the flu season, wildfire season, and allergy season). Thus, masks are not a normal part of life for people in the United States. However, in some places around the world, masks are normally worn due to heavy air pollution. For example, people living or working in the cities of China wear masks as a daily part of life. Every day, people wear masks to protect their lungs. Although pollution is decreasing due to policies set in place to fight pollution, it hasn't yet made masks an unnecessary part of everyday life.



Figure 7.2.*a*: Mask garbage. Image by Pixabay (Public Domain).

Pollution

7.2.1: Water Pollution

7.2.1.1: Water Pollutants and Their Sources

7.2.1.2: Water Treatment

7.2.1.3: Mitigating Water Pollution

7.2.2: Air Pollution

7.2.2.1: Ozone Depletion

7.2.2.2: Acid Deposition

7.2.3: Data Dive- Plastic in World Oceans

7.2.4: Review

Attribution

Modified by Rachel Schleiger (CC-BY-NC).

This page titled 7.2: Pollution is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

SECTION OVERVIEW

7.2.1: Water Pollution

Water pollution is the contamination of water by an excess amount of a substance that can cause harm to human beings and/or the ecosystem. Contaminants may come from from direct or indirect sources, or captured from air by rain. The level of water pollution depends on the abundance of the pollutant, the ecological impact of the pollutant, and the use of the water.

Attribution

Modified by Melissa Ha from Water Pollution from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

7.2.1.1: Water Pollutants and Their Sources

7.2.1.2: Water Treatment

7.2.1.3: Mitigating Water Pollution

This page titled 7.2.1: Water Pollution is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

7.2.1.1: Water Pollutants and Their Sources

Although natural processes such as volcanic eruptions or evaporation sometimes can cause water pollution, most pollution is derived from human, land-based activities. Pollutants can spread through different water reservoirs, as the water carrying them progresses through stages of the water cycle (figure 7.2.1.1.a). **Residence time** (the average time that a water molecule spends in a reservoir) is key to pollution problems because it affects pollution potential. Water in rivers has a relatively short residence time, so pollution usually is there only briefly. Of course, pollution in rivers may simply move to another reservoir, such as the ocean, where it can cause further problems. Groundwater is typically characterized by slow flow and longer residence time, which can make groundwater pollution particularly problematic. Finally, **pollution residence time** can be much greater than the water residence time because a pollutant may be taken up for a long time within the ecosystem or absorbed onto sediment.

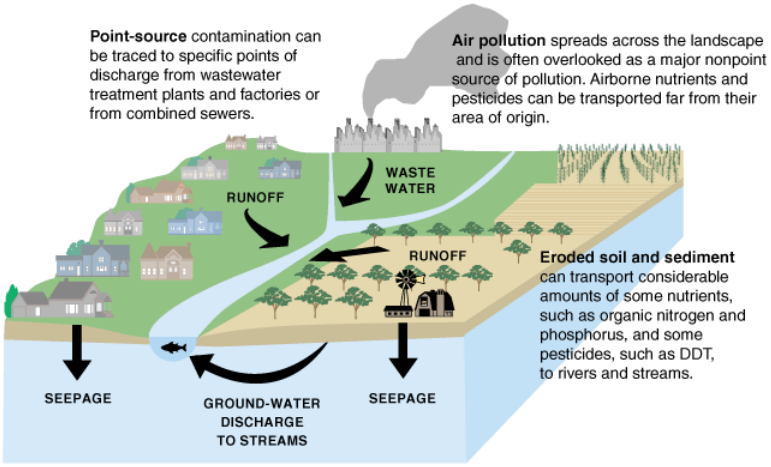


Figure 7.2.1.1.a: Sources of some water pollutants and movement of pollutants into different water reservoirs of the water cycle (such as rivers and aquifers). Point source contamination can be traced to specific points of discharge. Air pollution is an example of nonpoint source contamination because contaminants are throughout the air and landscape, ultimately entering the water from multiple locations. Eroded soil and sediment are another source of nonpoint source pollution. Image by USGS (public domain).

Water can be contaminated by various human activities or by existing natural features, like mineral-rich geologic formations. Agricultural activities, industrial operations, landfills, animal operations, and small and large scale sewage treatment processes, among many other things, all can potentially contribute to contamination. As water runs over the land or infiltrates into the ground, it dissolves material left behind by these potential contaminant sources. The risks and type of remediation for a contaminant depend on the type of chemicals present.

Point source pollution can be attributed to a single, definable origin. For example, animal factory farms (figure 7.2.1.1.b) raise a large number and high density of livestock such as cows, pigs, and chickens. Combined sewer systems that have a single set of underground pipes to collect both sewage and storm water runoff from streets for wastewater treatment can also be major point sources of pollutants. During heavy rain, storm water runoff may exceed sewer capacity, causing it to back up. This spills untreated sewage directly into surface waters (figure 7.2.1.1.c). Other examples include pipes from factories, waste disposal sites, storage tanks, and chemical spills.



Figure 7.2.1.1.*b*: Large animal farms are often referred to as concentrated feeding operations (CFOs). These farms are considered potential point sources of pollution because untreated animal waste may enter nearby waterbodies as untreated sewage. Credit: ehp.gov

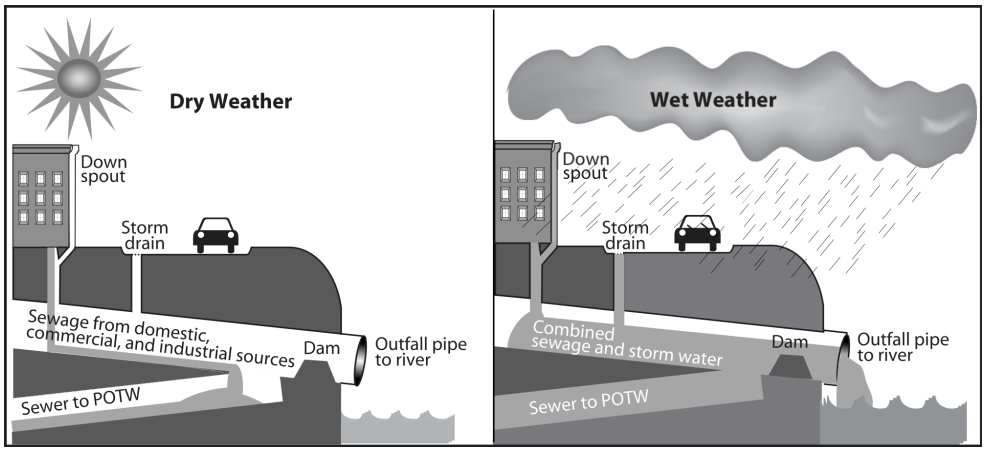


Figure 7.2.1.1.*c*: A combined sewer system is a possible major point source of water pollution during heavy rain due to overflow of untreated sewage. Sewage from domestic, commercial, and industrial sources enter the system through the down spout and storm water enters through the storm drain. During dry weather (and small storms), all sewage is handled by the publicly owned treatment works (POTW). A dam prevents raw sewage from flowing into a water body through the outfall pipe. During wet weather (large storms), the relief structure allows some of the combined sewage and storm water to be discharged untreated to an adjacent water body. Image by U.S. Environmental Protection Agency (public domain).

Nonpoint source pollution is from multiple dispersed sources. The whole of the contribution of pollutants is harmful, but the individual components may not reach harmful concentrations. Nonpoint sources of pollution include agricultural fields, cities, and abandoned mines. Rainfall runs over the land and through the ground, picking up pollutants from throughout the **watershed** (including areas of land and smaller streams that drain into a particular body of water). These pollutants might include herbicides, pesticides, and fertilizer from agricultural fields and lawns; oil, antifreeze, animal waste, and road salt from urban areas; and acid and toxic elements from abandoned mines. Then, this pollution is carried by runoff into surface water bodies and groundwater. Nonpoint source pollution, which is the leading cause of water pollution in the U.S., is usually much more difficult and expensive to control than point source pollution because of its low concentration, multiple sources, and much greater volume of water.

According to a 2016 report, there were 20,912 cases of impaired water bodies in the United States, which means they could neither support a healthy ecosystem nor meet water quality standards (table 7.2.1.1.*a*). Selected causes of impairment (water pollutants) are discussed below, categorized according to whether they arise from chemical, biological, or physical processes.

Table 7.2.1.1.*a***:** The leading causes of impaired bodies of water in the United States in 2016. Data from EPA (public domain).

Cause of Impairment	Number of Impaired Bodies of Water
Polychlorinated biphenyls (PCBs)	3,712
Pathogens	2,248
Nutrients	2,228
Mercury	2,138
Metals (other than mercury)	2,075
Cause unknown- impaired biota	1,852
Organic enrichment/oxygen depletion	1,281
Turbidity	1,175
Pesticides	795
Salinity/total dissolved solids/chlorides/sulfates	576
pH/acidity/caustic conditions	489
Sediment	453
Temperature	358
Total toxics	282
Algal growth	174
Cause unknown	159
Dioxins	136
Toxic organics	127
Toxic inorganics	99

Chemical Pollutants

Chemical pollution from agriculture, industry, cities, and mining threatens global water quality. Air pollutants from these activities can also enter bodies of water (and become water pollutants) through dry deposition, precipitation, and runoff. Some chemical pollutants have serious and well-known health effects, whereas many others have poorly known long-term health effects.

Any natural water contains dissolved chemicals, some of which are important human nutrients while others can be harmful to human health. The concentration of a water pollutant is commonly given in very small units such as parts per million (**ppm**) or even parts per billion (**ppb**). An arsenic concentration of 1 ppm means 1 part of arsenic per million parts of water. This is equivalent to one drop of arsenic in 50 liters of water. To give you a different perspective on appreciating small concentration units, converting 1 ppm to length units is 1 cm (0.4 in) in 10 km (6 miles) and converting 1 ppm to time units is 30 seconds in a year. **Total dissolved solids** (TDS) represent the total amount of dissolved material in water. Average TDS values for rainwater, river water, and seawater are about 4 ppm, 120 ppm, and 35,000 ppm, respectively.

Organic Pollutants

Organic pollutants include herbicides and pesticides, pharmaceuticals, fuel (such as oil spills), industrial solvents and cleansers, and synthetic hormones associated with pharmaceuticals. These synthetic hormones can act as **endocrine disruptors**. Many are **persistent organic pollutants** (POPs), which are long-lived in the environment, biomagnify through the food chain, and can be toxic. As previously mentioned, DDT (pesticide), dioxin (herbicide by-product), and PCBs (polychlorinated biphenyls, which were used as a liquid insulator in electric transformers), are all POPs.

An example of organic chemical contamination is the Love Canal, in Niagara Falls, New York (Figure 7.2.1.1.*d*). From 1942 to 1952, the Hooker Chemical Company disposed of over 21,000 tons of chemical waste, including chlorinated hydrocarbons, into a canal and covered it with a thin layer of clay. Chlorinated hydrocarbons are a large group of organic chemicals that have chlorine functional groups, most of which are toxic and carcinogenic to humans (DDT and PCBs are an examples.) The company sold the land to the New York School Board, who developed it into a neighborhood. After residents began to suffer from serious health ailments and pools of oily fluid started rising into residents' basements, the neighborhood had to be evacuated. This site became a U.S. Environmental Protection Agency **Superfund Site**, a site with federal funding and oversight to ensure its cleanup.



Figure 7.2.1.1.*d*: The Love Canal in Niagra Falls, New York. Image by Environmental Protection Agency (public domain).

Inorganic Pollutants

Inorganic pollutants (figure 7.2.1.1.*e*) include nutrients like nitrate (NO_3^-) and phosphate (PO_4^{3-}), heavy metals, chloride (Cl^-), and radioactive isotopes released from mining or nuclear accidents (such as cesium, iodine, uranium, and radon gas). Nutrients can be from geologic material, like phosphorus-rich rock, but are most often sourced from fertilizer and animal and human waste. Untreated sewage and agricultural runoff concentrate nitrogen and phosphorus which are essential for the growth of microorganisms. Nutrients like nitrate and phosphate in surface water can promote the growth of microbes, like blue-green algae (cyanobacteria), which in turn deplete dissolved oxygen (O_2) and produce toxins. This process is known as **eutrophication** (discussed below as well as previously in Biogeochemical Cycles, Threats to Biodiversity, and Industrial Agriculture).



Figure 7.2.1.1.*e*: This scientist demonstrates how to test water quality, and part of this involves measuring levels of inorganic pollutants. Image by Erin Gleeson (CC-BY).

Examples of heavy metals include arsenic, mercury, lead, cadmium, and chromium, and they can bioaccumulate and biomagnify through the food chain. Arsenic (As) enters the water supply naturally from weathering of arsenic-rich minerals and from human activities such as coal burning and smelting of metallic ores. The worst case of arsenic poisoning occurred in the densely populated impoverished country of Bangladesh, which had experienced 100,000s of deaths from diarrhea and cholera each year from drinking surface water contaminated with pathogens due to improper sewage treatment. In the 1970s the United Nations provided aid for millions of shallow water wells, which resulted in a dramatic drop in pathogenic diseases. Unfortunately, many of the wells produced water naturally rich in arsenic. Tragically, there are an estimated 77 million people (about half of the population) who inadvertently may have been exposed to toxic levels of arsenic in Bangladesh as a result. The World Health Organization has called it the largest mass poisoning of a population in history.

Mercury enters the water supply naturally from weathering of mercury-rich minerals and, like mercury, from human activities such as coal burning and metal processing. A famous mercury poisoning case in Minamata, Japan involved methylmercury-rich industrial discharge that caused high mercury levels in fish. People in the local fishing villages ate fish up to three times per day for over 30 years, which resulted in over 2,000 deaths. During that time the responsible company and national government did little to mitigate, help alleviate, or even acknowledge the problem.

Salt, typically sodium chloride, is a common inorganic contaminant. It can be introduced into groundwater from natural deposits or from anthropogenic sources like the salts applied to roads in the winter to keep ice from forming (figure 7.2.1.1.*f*). Salt contamination can also occur from **saltwater intrusion**, where cones of depression around fresh groundwater pumping near ocean coasts induce the encroachment of salt water into the freshwater body.



Figure 7.2.1.1.*f*: A snow plow and road salt application truck that is used to clear snow and ice from roadways for improving traffic and pedestrian safety. The use of salt to deice pavement can be harmful to aquatic life in urban streams. Image and caption (modified) from USGS (public domain).

The acidity or alkalinity of a body of water can also impact its quality. **pH** is a measure of the concentration of hydrogen ions (protons) in a solution, which determines how acidic or basic (alkaline) a solution is. Acidic solutions have a high hydrogen ion concentration and a pH less than 7, and basic solutions have a pH of greater than 7. The pH of freshwater usually ranges from 5 to 9, and salt water is slightly basic (pH = 8.2) in a healthy ecosystem. When conditions are too acidic, some aquatic animals cannot reproduce, and calcium carbonate structures (clams, snails, corals, etc.) dissolve. Acid deposition is further discussed later in this chapter, and ocean acidification is discussed in the next chapter.

Hard water contains abundant calcium and magnesium, which reduces its ability to develop soapsuds and enhances scale (calcium and magnesium carbonate minerals) formation on hot water equipment. Water softeners remove calcium and magnesium, which allows the water to lather easily and prevent the minerals from precipitating on surfaces (figure 7.2.1.1.*g*). Hard water develops naturally from the dissolution of calcium and magnesium carbonate minerals in soil; it does not have negative health effects in people.



Figure 7.2.1.1.*g*: Overly hard water (high calcium content) can cause film on glassware. Image by Liz Meimann/Iowa State Univ. (public domain).

Biological Pollutants

Pathogens (infectious microorganisms or viruses) enter water primarily from human and animal fecal waste due to inadequate sewage treatment. In many underdeveloped countries, sewage is discharged into local waters either untreated or after only rudimentary treatment. In developed countries untreated sewage discharge can occur from overflows of combined sewer systems,

poorly managed livestock factory farms, and leaky or broken sewage collection systems. Water with pathogens can be remediated by adding chlorine or ozone (O_3), by boiling, or by treating the sewage in the first place.

Physical Sources of Pollution

Trash, sediments, and thermal pollution arise from physical sources of pollution (figure 7.2.1.1.*h*). Trash was discussed extensively in Solid Waste Management. Excess sediments enter bodies of water when various land uses, such as mining, deforestation, and agriculture increases erosion. Sediments can carry toxins or excess nutrients with them, and they cloud the water (resulting in **turbidity**). Turbidity obstructs aquatic plants from accessing sufficient sunlight. Thermal pollution occurs when water temperature exceeds its natural range. Many power plants (such as coal, natural gas, nuclear, etc.) rely on water from the environment for cooling. This water is released back into bodies of water at a higher temperature than usual. High temperature disrupt aquatic organisms for several reasons; one is that warmer waters cannot hold as much dissolved oxygen (see below; figure 7.2.1.1.*i*). Dams can also raise water temperature to the detriment or the organisms that live there.



Figure 7.2.1.1.*h*: Water Pollution. Obvious water pollution in the form of floating debris; invisible water pollutants sometimes can be much more harmful than visible ones. Image by Stephen Codrington (CC-BY).

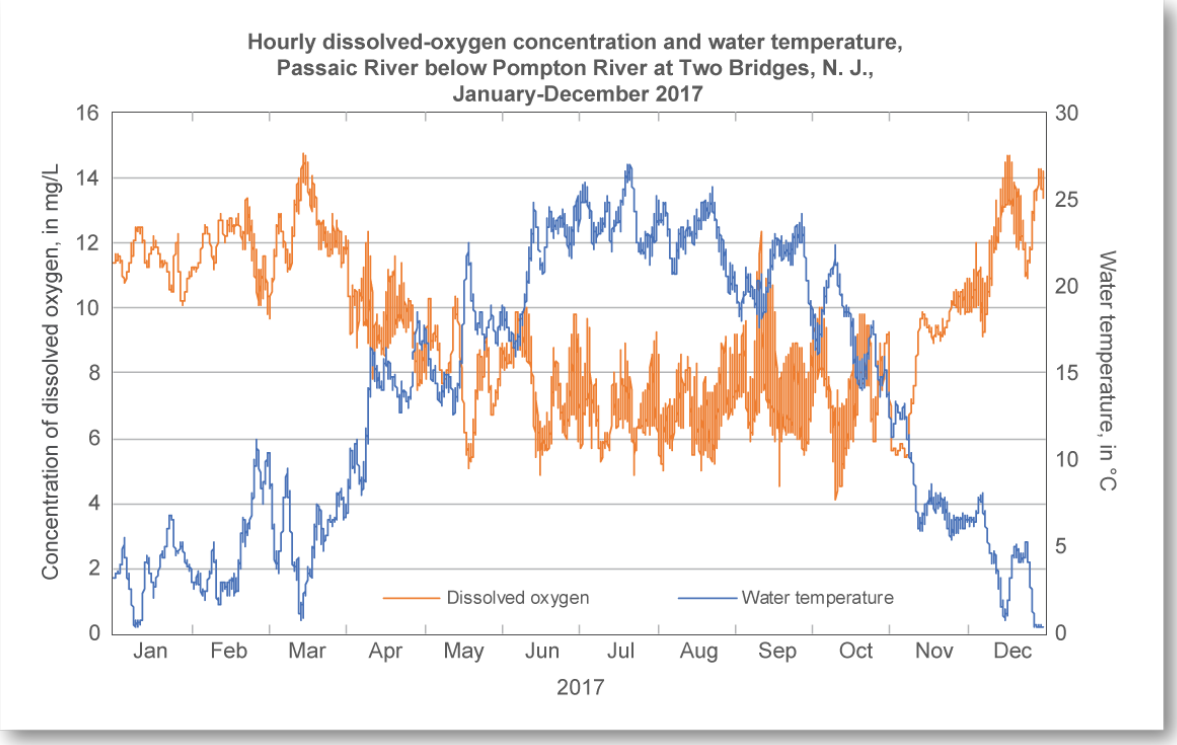


Figure 7.2.1.1.*i*: Water temperature affects dissolved oxygen concentrations in bodies of water. From January to December 2017, water temperature (blue) increases in summer months and then decreases. Dissolved oxygen levels (orange) dips in the summer when temperature increases. Image and caption (modified) from USGS (public domain).

Biochemical Oxygen Demand, Hypoxia, and Eutrophication

Oxygen-demanding waste is an extremely important pollutant to ecosystems. Most surface water in contact with the atmosphere has a small amount of dissolved oxygen, which is needed by aquatic organisms for cellular respiration. Decomposers, such as bacteria and fungi, also conduct cellular respiration and consume oxygen as they break down dead organic matter.

Too much decaying organic matter in water is a pollutant because it removes oxygen from water, which can kill fish, shellfish, and aquatic insects. The amount of oxygen used by **aerobic** (in the presence of oxygen) decomposition of organic matter is called **biochemical oxygen demand** (BOD). The major source of dead organic matter in many natural waters is sewage; grass and leaves are smaller sources.

An unpolluted water body with respect to BOD is a turbulent river that flows through a natural forest. Turbulence continually brings water in contact with the atmosphere where the dissolved oxygen content is restored. The dissolved oxygen content in such a river ranges from 10 to 14 ppm. When biological oxygen demand is low, clean-water fish such as trout thrive (figure 7.2.1.1.*j*).



Figure 7.2.1.1.j: Left: Fish like rainbow trout thrive under low biological oxygen demand and ample levels of dissolved oxygen. Right: Fish like carp thrive under high biological oxygen demand and low dissolved oxygen levels. Left image by Engbretson Eric/U.S. Fish and Wildlife Service (public domain). Right image by X posid (public domain).

A polluted water body with respect to BOD is a stagnant, deep lake in an urban setting with a combined sewer system. This system favors a high input of dead organic carbon from sewage overflows and limited chance for water circulation and contact with the atmosphere. In such a lake, the dissolved oxygen content is ≤ 5 ppm. Biological oxygen demand is high, and fish tolerant to low oxygen levels, such as carp and catfish dominate (figure 7.2.1.1.k).

Excessive nutrients, particularly nitrogen (N) and phosphorous (P), are pollutants closely related to oxygen-demanding waste. Plants and algae require 15-20 nutrients for growth, most of which are plentiful in water. Nitrogen and phosphorus are called **limiting nutrients**, however, because they usually are present in water at low concentrations and therefore restrict the total amount of plant growth. This explains why nitrogen and phosphorus are major ingredients in most fertilizer.

High concentrations of limiting nutrients, particularly nitrogen (N) and phosphorous (P), from human sources (mostly agricultural and urban runoff including fertilizer, sewage, and phosphorus-based detergent) can cause cultural eutrophication, which leads to the rapid growth of aquatic producers, particularly algae. Thick mats of floating algae or rooted plants lead to a form of water pollution that damages the ecosystem by clogging fish gills and blocking sunlight (figure 7.2.1.1.k). A small percentage of algal species produce toxins that can kill animals, including humans. Exponential growths of these algae are called **toxic algal blooms**.



Figure 7.2.1.1.*k*: A cyanobacteria (blue-green algae) bloom in Lake Erie resulting from fertilizer pollution. Image by NASA (public domain).

When the prolific algal layer dies, it becomes oxygen-demanding waste, which can create very low oxygen concentrations in the water (< 2 ppm), a condition called **hypoxia.** This results in a dead zone because it causes death from asphyxiation to organisms that are unable to leave that environment (figure 7.2.1.1.*l*). Approximately 50% of lakes in North America, Europe, and Asia are negatively impacted by eutrophication. Eutrophication and hypoxia are difficult to combat because they are caused primarily by nonpoint source pollution, which is difficult to regulate, and nitrogen and phosphorus are difficult to remove from wastewater.

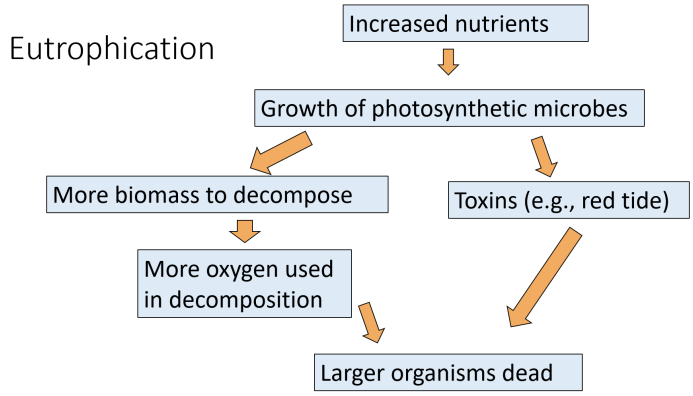


Figure 7.2.1.1.*l*: The process of eutrophication. During eutrophication, increased nutrients causes the growth of photosynthetic microbes resulting in more biomass to decompose. More oxygen is used in decomposition, resulting in the death of larger organisms. The growth of photosynthetic microbes also leads to toxins (for example, those produced during red tides), which also leads to the death of larger organisms.

Attributions

Modified by Melissa Ha from the following sources:

- Water Contamination from *An Introduction to Geology* by Johnson et al. (licensed under CC-BY-NC-SA)
- Water Pollution from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- The Industrialization of Nature: A Modern History (1500 to the present) from *Sustainability: A Comprehensive Foundation* by Tom Theis and Jonathan Tomkin, Editors. Download for free at CNX. (licensed under CC-BY)

This page titled 7.2.1.1: Water Pollutants and Their Sources is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **20.1.1: Water Pollutants and Their Sources** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

7.2.1.2: Water Treatment

The most deadly form of water pollution, pathogenic microorganisms that cause waterborne diseases, killed 485,000 million people (mostly in developing countries) in 2017. The best strategy for addressing this problem is proper sewage (wastewater) treatment. Untreated sewage is not only a major cause of pathogenic diseases, but also a major source of other pollutants, including oxygen-demanding waste, excess nutrients, and toxic heavy metals. Wastewater treatment is done at a sewage treatment plant in urban areas and through a septic tank system in rural areas.

The wastes generated by some 80% of U.S. households are collected in sewer systems. Each day, the United States processes about 34 billion gallons of wastewater (99.9% of this is water). In 2004, an Environmental Protection Agency (EPA) report estimated that 860 billion gallons of untreated wastewater (raw sewage) escaped into rivers, streams, and the ocean, mainly from combined sewage systems (see Water Pollutants and Their Sources). As of 2020, the EPA has assessed nearly 96% of the systems responsible for wastewater overflow and addressed them with a civil complaint, enforcement order, or permit requirement. However, in 2017, the EPA estimated that were still 40,000 or more wastewater overflow events, which released raw sewage into bodies of water.

To reduce water pollution problems, separate sewer systems (where street runoff goes to rivers and only wastewater goes to a wastewater treatment plant) are much better than combined sewer systems, which can overflow and release untreated sewage into surface waters during heavy rain. Some cities such as Chicago, Illinois have constructed large underground caverns and also use abandoned rock quarries to hold storm sewer overflow. After the rain stops, the stored water goes to the sewage treatment plant for processing.

Wastewater Treatment

The sewage that is properly directed to a wastewater treatment plant undergoes several steps: pretreatment, primary treatment, tertiary treatment, and disinfection and discharge (figure 7.2.1.2.*a*). **Pretreatment** involves the screening and removal of sand and gravel. **Primary treatment** involves settling or floatation to remove organic solids, fat, and grease. The undissolved solids in raw sewage to settle out of suspension forming **sludge**. Such **primary treatment** removes only one-third of the biological oxygen demand (BOD) and virtually none of the dissolved minerals. Attempts to use digested sludge as a fertilizer have been hampered by its frequent contamination by toxic chemicals derived from industrial wastes.

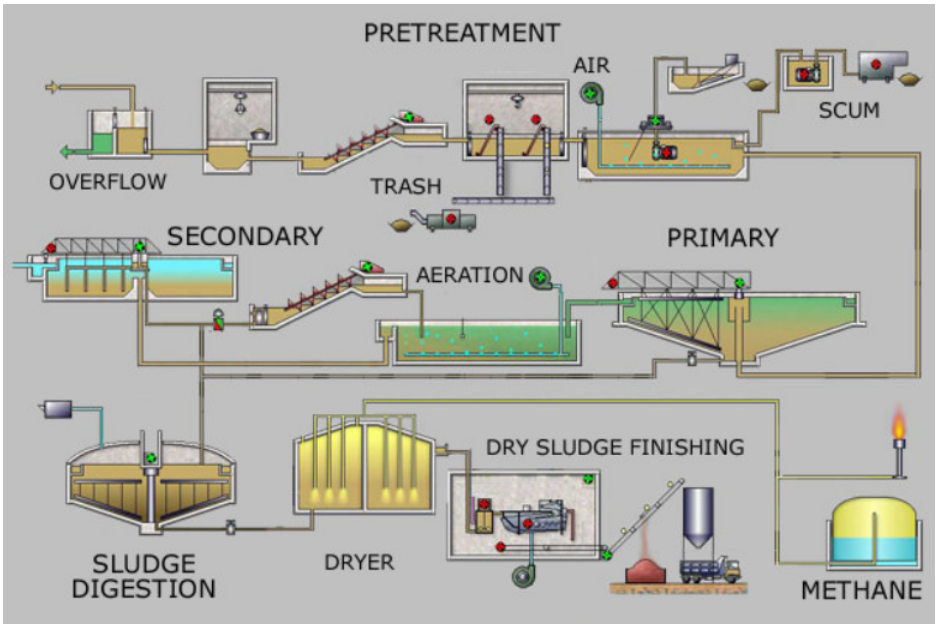


Figure 7.2.1.2.*a*: Steps at a wastewater treatment plant. The numerous processing steps at a conventional sewage treatment plant include pretreatment (screening and removal of sand and gravel), primary treatment (settling or floatation to remove organic solids, fat, and grease), secondary treatment (aerobic bacterial decomposition of organic solids), tertiary treatment (bacterial decomposition of nutrients and filtration), disinfection (chemical or UV treatment to kill bacteria), and either discharge to surface waters or reuse. Only pretreatment, primary treatment, secondary treatment, and sludge digestion are shown here. Sewage initially enters pretreatment, though some may escape due to overflow. Screening removes larger components, which enter the trash. Air is added, the top layer (scum) is removed, and sludge (solids that settle to the bottom) are removed during primary treatment. Further aeration is required for secondary treatment to supply oxygen to the aerobic bacteria that break down organic components in the fluid. More sludge is removed. The sludge removed during primary and secondary treatment undergo sludge digestion. This involves a dry, dry sludge finishing, and methane production. Image by Leonard G. (CC-BY-SA).

Next, the organic solids in the effluent are decomposed by **aerobic** (oxygen-requiring) bacteria in **secondary treatment**. Here the effluent is brought in contact with oxygen. Aerobic microorganisms break down much of the organic matter to harmless substances such as carbon dioxide. Primary and secondary treatment together can remove up to 90% of the BOD. After chlorination to remove its content of bacteria, the effluent from secondary treatment is returned to the local surface water.

The concentrated organic solid produced during primary and secondary treatment (sludge) is treated in a variety of ways including landfill disposal, incineration, use as fertilizer, and **anaerobic** bacterial decomposition, which is done in the absence of oxygen. Anaerobic decomposition of sludge produces methane gas, which can be used as an energy source.

Although the combination of primary and secondary treatment removes most of the organic matter in sewage, lowering BOD, most of the nitrogen and phosphorus in sewage remains in the effluent (liquid portion of the sewage). In **tertiary treatment**, bacteria are used to remove the remaining nutrients, and filtration occurs. Additionally, several techniques are available to remove dissolved salts from sewage effluent, but all are quite expensive.

During **disinfection**, further treatment involves disinfection with chlorine, ozone, ultraviolet light, or bleach to kill most microbes. Finally, the treated water is **discharged** to surface waters (usually a local river) or **reused** for some other purpose, such as irrigation, habitat preservation, and artificial groundwater recharge

Septic Tank System

A **septic tank system** is an individual sewage treatment system for homes in typically rural settings. The basic components of a septic tank system (figure 7.2.1.2.*b*) include a sewer line (inlet) from the house, a septic tank, and the drain field. The **septic tank** is a large container where sludge settles to the bottom, where microorganisms decompose the organic solids anaerobically. The **drain field** is a network of perforated pipes where the clarified water seeps into the soil and is further purified by bacteria. Water pollution problems occur if the septic tank malfunctions, which usually occurs when a system is established in the wrong type of soil or maintained poorly.

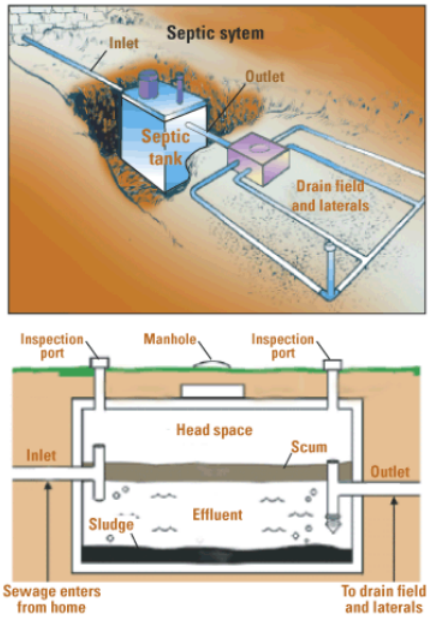


Figure 7.2.1.2.*b*: A septic tank system for sewage treatment. The inlet transports sewer to the septic tank (top). Inside the septic tank (bottom), sludge settles to the bottom of the liquid effluent. Anaerobic bacteria digest the sludge. Scum floats on the top of the effluent, and the air above it is called the headspace. Two inspection ports extend above the septic tank, which is buried underground. It is also accessible through a manhole. Effluent leaves the outlet into the drain field, which contain pipes (laterals) with small holes in them. The effluent is released and can infiltrate into the ground and be further purified by bacteria. Image by United States Geological Survey (public domain).

Addressing Water Treatment Globally

For many developing countries, financial aid is necessary to build adequate wastewater treatment facilities. The World Health Organization estimates an estimated cost savings at least $5 for every $1 invested in clean water delivery and sanitation. The cost savings are from health care savings, gains in work and school productivity, and prevented deaths. Simple and inexpensive techniques for treating water at home include chlorination, filters, and solar disinfection. Another alternative is to use constructed wetlands technology (marshes built to treat contaminated water), which is simpler and cheaper than a conventional sewage treatment plant.

Attributions

Modified by Melissa Ha from the following sources:

- Water Treatment from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Sewage Treatment from *Biology* by John W. Kimball (licensed under CC-BY)
- Former National Compliance Initiative: Keeping Raw Sewage and Contaminated Stormwater Out of Our Nation's Waters. 2020. EPA. Accessed 01-18-2021. (public domain)

This page titled 7.2.1.2: Water Treatment is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **20.1.2: Water Treatment** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

7.2.1.3: Mitigating Water Pollution

Water pollution can be mitigated through regulation, bioremediation, and watershed management.

Regulation

During the early 1900s, rapid industrialization in the U.S. resulted in widespread water pollution due to free discharge of waste into surface waters. The Cuyahoga River in northeast Ohio caught fire numerous times, including a famous fire in 1969 that caught the nation's attention (figure 7.2.1.3.a). In 1972 Congress passed one of the most important environmental laws in U.S. history, the Federal Water Pollution Control Act, which is more commonly called the **Clean Water Act** (CWA). The purpose of the Clean Water Act and later amendments is to maintain and restore water quality, or in simpler terms to make our water swimmable and fishable. It became illegal to dump pollution into surface water unless there was formal permission. The CWA regulates pollution from a single source such as that from industry or sewage treatment plants by setting pollution standards (maximum levels of each pollutant that can be in bodies of water or released at a time). United States water quality improved significantly as a result, but as there is still more work to be done.



Figure 7.2.1.3.a: Cuyahoga River on fire in 1969. Image from USEPA Environmental-Protection-Agency (public domain).

Remediation

Remediation is the act of cleaning contamination. Biological remediation (**bioremediation**) is a waste management technique that involves the use of organisms such as plants, bacteria, and fungi to remove or neutralize pollutants from a contaminated site. According to the United States EPA, bioremediation is a "treatment that uses naturally occurring organisms to break down hazardous substances into less toxic or nontoxic substances". This type of remediation is usually used on organic chemicals but also works on reducing or oxidizing inorganic chemicals like nitrate. **Phytoremediation** is a type of bioremediation that uses plants to absorb the chemicals over time.

Bioremediation is widely used to treat human sewage and has also been used to remove agricultural chemicals (pesticides and fertilizers) that leach from soil into groundwater. Certain toxic metals, such as selenium and arsenic compounds, can also be removed from water by bioremediation. Mercury is an example of a toxic metal that can be removed from an environment by bioremediation. Several species of bacteria can carry out the biotransformation of toxic mercury into nontoxic forms. These bacteria, such as *Pseudomonas aeruginosa*, can convert a charged form of mercury (Hg^{2+}) to an uncharged form (Hg), which is less toxic to humans.

Probably one of the most useful and interesting examples of the use of prokaryotes for bioremediation purposes is the cleanup of oil spills. To clean up these spills, bioremediation is promoted by adding inorganic nutrients that help bacteria already present in the environment to grow. Hydrocarbon-degrading bacteria feed on the hydrocarbons in the oil droplet, breaking them into inorganic compounds. Some species, such as *Alcanivorax borkumensis*, produce surfactants that break the oil into droplets, making it more accessible to the bacteria that degrade the oil. In the case of oil spills in the ocean, ongoing, natural bioremediation tends to occur, inasmuch as there are oil-consuming bacteria in the ocean prior to the spill. Under ideal conditions, it has been reported that up to 80 percent of the nonvolatile components (those that do not readily evaporate) in oil can be degraded within one year of the spill.

Researchers have genetically engineered other bacteria to consume petroleum products; indeed, the first patent application for a bioremediation application in the U.S. was for a genetically modified oil-eating bacterium.



(a) (b)

Figure 7.2.1.3.a: (a) Cleaning up oil after the Exxon Valdez spill in Alaska (1989), the workers hosed oil from beaches and then used a floating boom to corral the oil, which was finally skimmed from the water surface. Some species of bacteria are able to solubilize and degrade the oil. (b) One of the most catastrophic consequences of oil spills is the damage to fauna. (credit a: modification of work by NOAA; credit b: modification of work by GOLUBENKOV, NGO: Saving Taman)

There are a number of cost/efficiency advantages to bioremediation, which can be employed in areas that are inaccessible without excavation. For example, hydrocarbon spills (specifically, oil spills) or certain chlorinated solvents may contaminate groundwater, which can be easier to treat using bioremediation than more conventional approaches. This is typically much less expensive than excavation followed by disposal elsewhere, incineration, or other off-site treatment strategies. It also reduces or eliminates the need for "pump and treat", a practice common at sites where hydrocarbons have contaminated clean groundwater. Using microorganisms for bioremediation of hydrocarbons also has the advantage of breaking down contaminants at the molecular level, as opposed to simply chemically dispersing the contaminant.

Chemical remediation uses the introduction of chemicals to remove the contaminant or make it less harmful. One example is reactive barriers, a permeable wall in the ground or at a discharge point that chemically reacts with contaminants in the water. Reactive barriers made of limestone can increase the pH of acid mine drainage, making the water less acidic and more basic, which removes dissolved contaminants by precipitation into a solid form. **Physical remediation** consists of removing the contaminated water and either treating it (aka pump and treat) with filtration or disposing of it. All of these options are technically complex, expensive, and difficult, with physical remediation typically being the most costly.

Watershed Management

Watershed management involves reducing chemicals applied to land in the watershed (which will drain into a body of water) and runoff of those chemicals (figure 7.2.1.3.$b-c$). This strategy is more effective for nonpoint source pollution than setting pollution standards (as the CWA does) because it does not require each source of pollution be identified.



Figure 7.2.1.3.*b*: While some watersheds are relatively small, others encompass thousands of square miles and may contain streams, rivers, lakes, reservoirs, and underlying groundwater that are hundreds of miles inland. Shown here is an aerial view of Drakes Bay, part of California's Tomales-Drake watershed. Image and caption (modified) from Brian Cluer, NOAA Fisheries West Coast Region, California Coastal Office (public domain).

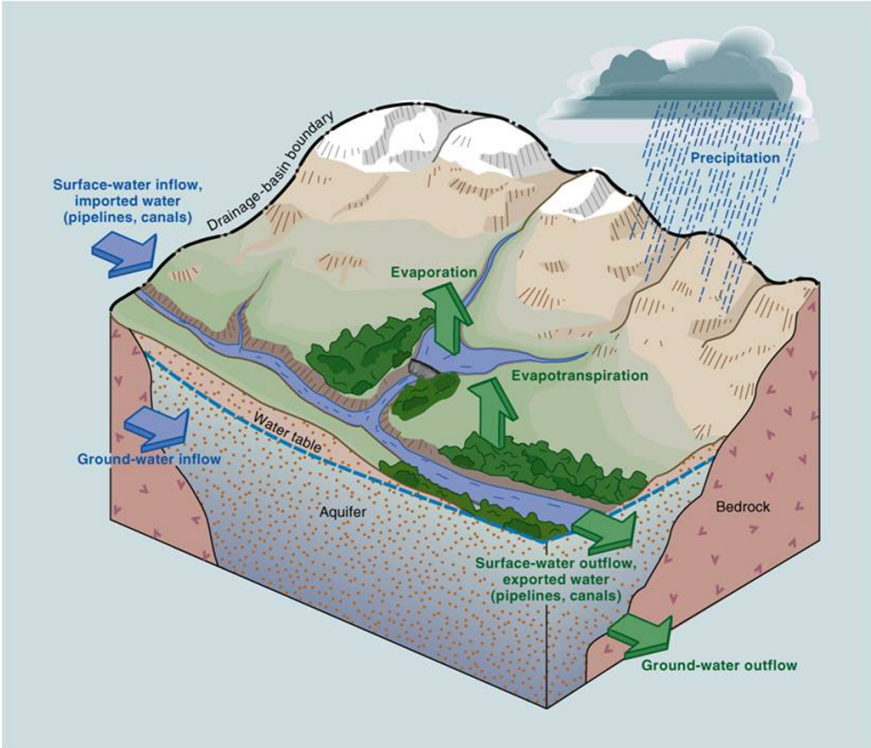


Figure 7.2.1.3.*c*: Diagram of a watershed, the area that drains into a body of water. The drainage-basin boundary defines the edges of the watershed. Precipitation falls on the watershed. Some reenters the atmosphere through evaporation or evapotranspiration. The water collects into smaller streams, which drain into a larger river. Additionally, surface water can enter the watershed (surface-water inflow) as water imported through pipelines or canals. Surface water exits the watershed (surface-water outflow) as water exported through pipelines canals. Similarly, ground-water inflow enters the system, and ground-water outflow exits the system. Ground water is stored in the aquifer underground, and the water table marks upper level of groundwater in the aquifer. Image from Healy, R.W., Winter, T.C., LaBaugh, J.W., and Franke, O.L., 2007, : Water budgets: Foundations for effective water-resources and environmental management: U.S. Geological Survey Circular 1308.

Maintaining or restoring **riparian areas** (riparian zones) is key to watershed management (figure 7.2.1.3.*d*). These are areas of land close enough to a body of water to be affected by that body of water, for example, the lush region of vegetation surrounding a river. Riparian areas provide many ecosystem services that promote water quality and limit pollution. The vegetation absorbs nutrients that could otherwise cause eutrophication. It also provide shade, keeping the water cool and increasing its capacity to hold dissolved oxygen. The roots of the vegetation cling to soil, combating erosion. Additionally, roots and low-lying plants slow the flow of run-off, promoting infiltration. High infiltration rates have several benefits: (1) less runoff is available to bring pollution to the river, lake, or bay, (2) pollutants will be filtered from the runoff as it soaks into the ground, and (3) aquifers are replenished.



Figure 7.2.1.3.d: Riparian areas reduce water pollution by absorbing nutrients, keeping water cool, limiting erosion, and promoting infiltration. Image by BLM (public domain).

Watershed management plans typically leave undisturbed riparian vegetation directly adjacent to a body of water (figure 7.2.1.3.e). Next (moving away from the water), could be managed forest followed by agriculture. The most intensive agriculture should be farthest from the body of water, particularly when pesticides and fertilizers are in use.

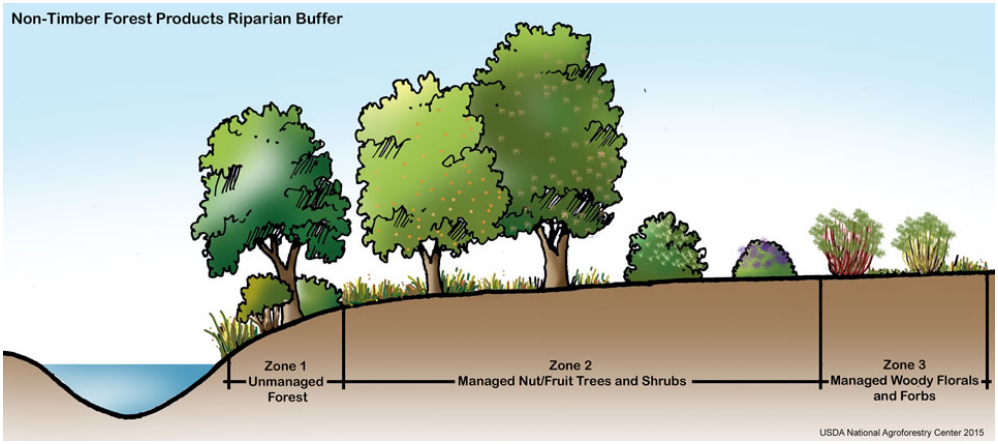


Figure 7.2.1.3.e: Although riparian forest buffers are often designed primarily for water quality benefits, these practices can also include woody species that provide products such as nuts, fruit, and decorative woody florals. These are examples of non-timber forest products. Zone 1, which is closest to the river, is unmanaged forest. Zone 2 contains managed nut or fruit trees and shrubs. Zone 3 contains managed woody florals and forbs. (Forbs are non-woody plants excluding grasses.) Image and caption by USDA National Agroforestry Center (public domain).

In urban areas can also be carefully structured to limit water pollution. Rain gardens next to buildings are areas with soil and vegetation that promote infiltration. Permeable pavement, which allows water to pass through it, also promotes infiltration, and reduces runoff, creating fewer opportunities for the water to acquire pollutants from washing over a dirty street.

Attribution

Modified by Melissa Ha from the following sources:

- Water Contamination from *An Introduction to Geology* by Johnson et al. (licensed under CC-BY-NC-SA)
- Water Treatment and Bioremediation from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 7.2.1.3: Mitigating Water Pollution is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **20.1.3: Mitigating Water Pollution** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC-SA 4.0.

SECTION OVERVIEW

7.2.2: Air Pollution

Air pollution occurs in many forms but can generally be thought of as gaseous and particulate contaminants that are present in the earth's atmosphere. Chemicals discharged into the air that have a direct impact on the environment are called **primary pollutants**. These primary pollutants sometimes react with other chemicals in the air to produce **secondary pollutants**.

Air pollution is typically separated into two categories: outdoor air pollution and indoor air pollution. **Outdoor air pollution** involves exposures that take place outside of the built environment. Examples include fine particles produced by the burning of coal, noxious gases such as sulfur dioxide, nitrogen oxides and carbon monoxide; ground-level ozone and tobacco smoke. **Indoor air pollution** involves exposures to particulates, carbon oxides, and other pollutants carried by indoor air or dust. Examples include household products and chemicals, out-gassing of building materials, allergens (cockroach and mouse dropping, mold, pollen), and tobacco smoke.

This video summarizes the causes, consequences, and solutions to air pollution.



Sources of Air Pollution

A **stationary source** of air pollution refers to an emission source that does not move, also known as a point source. Stationary sources include factories, power plants, and dry cleaners. The term **area source** is used to describe many small sources of air pollution located together whose individual emissions may be below thresholds of concern, but whose collective emissions can be significant. Residential wood burners are a good example of a small source, but when combined with many other small sources, they can contribute to local and regional air pollution levels. Area sources can also be thought of as non-point sources, such as construction of housing developments, dry lake beds, and landfills.

A **mobile source** of air pollution refers to a source that is capable of moving under its own power. In general, mobile sources imply "on-road" transportation, which includes vehicles such as cars, sport utility vehicles, and buses. In addition, there is also a "non-road" or "off-road" category that includes gas-powered lawn tools and mowers, farm and construction equipment, recreational vehicles, boats, planes, and trains.

Agricultural sources arise from operations that raise animals and grow crops, which can generate emissions of gases and particulate matter. For example, animals confined to a barn or restricted area produce large amounts of manure. Manure emits various gases, particularly ammonia into the air. This ammonia can be emitted from the animal houses, manure storage areas, or from the land after the manure is applied. In crop production, the misapplication of fertilizers, herbicides, and pesticides can potentially result in aerial drift of these materials and harm may be caused.

Unlike the above mentioned sources of air pollution, air pollution caused by **natural sources** is not caused by people or their activities. An erupting volcano emits particulate matter and gases, forest and prairie fires can emit large quantities of pollutants (figure 7.2.2.a), dust storms can create large amounts of particulate matter, and plants and trees naturally emit volatile organic

compounds which can form aerosols that can cause a natural blue haze. Wild animals in their natural habitat are also considered natural sources of “pollution”.



Figure 7.2.2.*a*: Smoke from the Camp fire in November of 2018 covers much of Northern California and extends over the Pacific Ocean. Image by NASA Earth Observatory/Lauren Dauphin (public domain).

Criteria Air Pollutants

The Environmental Protection Agency has identified six common air pollutants called **criteria air pollutants**. They are particulate matter, ground-level ozone, carbon monoxide, sulfur oxides, nitrogen oxides, and lead. These pollutants can harm health and the environment, and cause property damage. Of the six pollutants, particle pollution and ground-level ozone are the most widespread health threats. The U.S. Environmental Protection Agency (EPA) regulates them by developing criteria based on considerations of human and environmental health.

1. **Ground-level ozone** is not emitted directly into the air, but is created by chemical reactions between nitrogen oxides (NO_X) and the gaseous oxygen (O_2) in the atmosphere react in the presence of sunlight. It also forms when nitrogen oxides and volatile organic compounds (VOCs) react. Emissions from industrial facilities and electric utilities, motor vehicle exhaust, gasoline vapors, and chemical solvents are some of the major sources of NO_X and VOCs. Breathing ozone can trigger a variety of health problems, particularly for children, the elderly, and people of all ages who have lung diseases such as asthma. Ground level ozone can also have harmful effects on sensitive vegetation and ecosystems. (Ground-level ozone should not be confused with the ozone layer, which is high in the atmosphere and protects Earth from ultraviolet light; ground-level ozone provides no such protection).
2. **Particulate matter**, also known as particle pollution, is a complex mixture of extremely small particles and liquid droplets. Particle pollution is made up of a number of components, including acids (such as nitrates and sulfates), organic chemicals, metals, and soil or dust particles. Examples of human-generated particulate matter are from combustion (such as burning wood or coal) and mining (which releases dust). Wildfires or wind erosion are natural sources of particulate matter. The size of particles is directly linked to their potential for causing health problems. EPA is concerned about particles that are 10 micrometers in diameter or smaller because those are the particles that generally pass through the throat and nose and enter the lungs. Furthermore, particles that are 2.5 micrometers in diameter or smaller can pass from the lungs into the bloodstream, through which they are transported to various organs and cause serious health effects.
3. **Carbon monoxide** (CO) is a colorless, odorless gas emitted from combustion processes. Nationally and, particularly in urban areas, the majority of CO emissions to ambient air come from vehicles. However, volcanic eruptions and wildfires also release carbon monoxide. Carbon monoxide can cause harmful health effects by reducing oxygen delivery to the body’s organs (like the heart and brain) and tissues. At extremely high levels, CO can cause death.
4. **Nitrogen oxides** (NO_X) are a group of highly reactive gasses, which includes nitrogen monoxide (NO) and nitrogen dioxide (NO_2). EPA’s National Ambient Air Quality Standard uses NO_2 as the indicator for the larger group of nitrogen oxides. Nitrogen oxides forms quickly from emissions from cars, trucks and buses, power plants, and off-road equipment. Additionally, nitrogen oxides are released naturally by certain bacteria. In addition to contributing to the formation of ground-level ozone, and fine particle pollution, nitrogen oxides react in the atmosphere to form nitric acid (HNO_3) and nitrous acid (HNO_2), which are components of Acid Deposition. Nitrogen oxides is linked with a number of adverse effects on the respiratory system.

5. **Sulfur dioxide** (SO_2) is mainly released from fossil fuel combustion at power plants (73%) and other industrial facilities (20%). Smaller sources of SO_2 emissions include industrial processes such as extracting metal from ore, and the burning of high sulfur containing fuels by locomotives, large ships, and non-road equipment. Volcanoes are a natural source of SO_2. Sulfur dioxide is linked with a number of adverse effects on the respiratory system. It reacts in the atmosphere to form sulfuric acid (H_2SO_4), a component of Acid Deposition.
6. **Lead** is a metal found naturally in the environment as well as in manufactured products. The major sources of lead emissions have historically been from fuels in on-road motor vehicles (such as cars and trucks) and industrial sources. As a result of regulatory efforts in the U.S. to remove lead from on-road motor vehicle gasoline, emissions of lead from the transportation sector dramatically declined by 95 percent between 1980 and 1999, and levels of lead in the air decreased by 94 percent between 1980 and 1999. Today, the highest levels of lead in air are usually found near lead smelters. The major sources of lead emissions to the air today are ore and metals processing and piston-engine aircraft operating on leaded aviation gasoline. Volcanoes and erosion are natural sources of lead.

Indoor Air Pollution (Major concerns in developed countries)

Most people spend approximately 90 percent of their time indoors. However, the indoor air we breathe in homes and other buildings can be more polluted than outdoor air and can increase the risk of illness. There are many sources of indoor air pollution in homes. They include biological contaminants such as bacteria, molds and pollen, burning of fuels and environmental tobacco smoke, building materials and furnishings, household products, central heating and cooling systems, and outdoor sources. Outdoor air pollution can enter buildings and become a source of indoor air pollution.

Sick building syndrome is a term used to describe situations in which building occupants have health symptoms that are associated only with spending time in that building. Causes of sick building syndrome are believed to include inadequate ventilation, indoor air pollution, and biological contaminants. Usually indoor air quality problems only cause discomfort. Most people feel better as soon as they remove the source of the pollution. Making sure that your building is well-ventilated and getting rid of pollutants can improve the quality of your indoor air.

Secondhand Smoke (Environmental Tobacco Smoke)

Secondhand smoke is the combination of smoke that comes from a cigarette and smoke breathed out by a smoker. When a non-smoker is around someone smoking, they breathe in secondhand smoke.

Secondhand smoke is dangerous to anyone who breathes it in. There is no safe amount of secondhand smoke. It contains over 7,000 harmful chemicals, at least 250 of which are known to damage human health. It can also stay in the air for several hours after somebody smokes. Even breathing secondhand smoke for a short amount of time can hurt your body.

Over time, secondhand smoke can cause serious health issues in non-smokers. The only way to fully protect non-smokers from the dangers of secondhand smoke is to not allow smoking indoors. Separating smokers from nonsmokers (like “no smoking” sections in restaurants), cleaning the air, and airing out buildings does not completely get rid of secondhand smoke.

Source: Smokefree.gov

Addressing Air Pollution

A combination of regulatory, economic, and technological strategies can reduce air pollution. The 1970 **Clean Air Act** is a key federal regulation that has successfully reduced emissions of air pollutants such as lead and sulfur dioxide. The Environmental Protection Agency sets an air quality standard for large emitters of criteria air pollutants, which is the maximum allowable amount of the pollutant that can be released. Each state then develops a plan to comply with these standards. Note that hundreds of air pollutants beyond the six criteria air pollutants are regulated under the Clean Air Act.

Economic incentives make it financially beneficial for individuals, institutions, or companies to use technologies or otherwise take action that limits air pollution. For examples, **green taxes** increase the cost of an environmentally harmful action. For example, Canada levies a green tax on fuel-inefficient vehicles. **Subsidies**, on the other hand reduce the cost of environmentally beneficial choices. For example, the federal solar tax credit in the United States allows homeowners to claim a percentage of the cost of

installing solar panels when their filing taxes This amount would thus reduce taxes owed or be issued as a tax refund. (Increased usage of solar electricity ultimately reduces the need for coal and natural gas power plants, which generate air pollution.)

Tradable permits (**cap-and trade**) is another economic incentive. First, the total allowable amount of emissions for an entity (such as a state or country) is set. Next, emission credits are distributed to potential polluters. Industries that reduce emissions can sell their credits, while those with high emissions may be required. For example, large emitters of greenhouse gases in California must by carbon pollution permits. The purchase of these permits funds the California Climate Credit, which Californians receive as a reduction to their April and October electricity bills.

Finally, a variety of technologies reduce air pollution. For example, air filters and proper ventilation are essential for limiting indoor air pollution. Advances in public transportation (and policies that promote them), and energy-efficient vehicles further reduce air pollution. Electric vehicles also reduce air pollution (figure 7.2.2.*b*), particularly if the the electricity that supplies them comes from a clean, renewable source, such as solar panels. **Catalytic converters** make emissions from standard vehicles less harmful by facilitating chemical reactions. For example, they catalyze reactions that convert nitrogen oxides (NO_X) to nitrogen gas (N_2) and carbon monoxide (CO) to carbon dioxide (CO_2). (While carbon dioxide is a greenhouse gas, it is not directly poisonous as carbon monoxide is.)

Other technologies limit emissions from industry. For example, **smokestack scrubbers** remove some pollutants from power plant emissions before they are released and have reduced sulfur dioxide emissions from burning coal. **Electrostatic precipitators** negatively charge pollutants and remove charged pollutants with a positive electrode (figure 7.2.2.*c*). While electrostatic precipitators were originally developed for industry, smaller versions can be used as air purifiers in homes or businesses.



Figure 7.2.2.*b*: Mass transportation, especially those that use clean energy, reduces air pollution. Image by Mtattrain (CC-BY-SA).

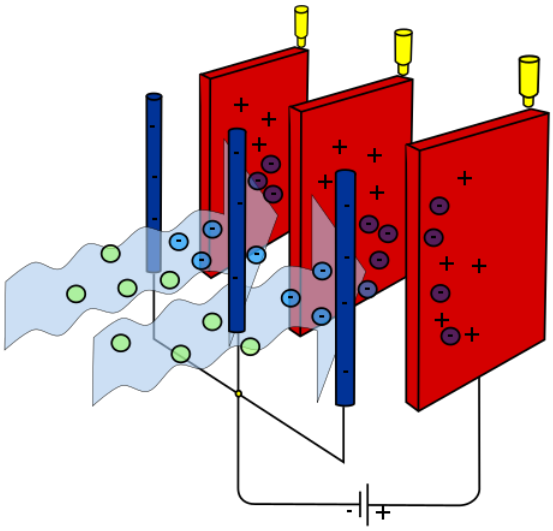


Figure 7.2.2.*c*: An electrostatic precipitator. Blue arrows represent the flow of air past metal rods, which negatively charge the pollutants. These are then drawn to positive electrodes (red, flat metal plates), removing the pollutants form the air. Image by Evan Mason (CC-BY-SA).

Attribution

Modified by Melissa Ha from the following sources:

- Atmospheric Pollution from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Basic Information about NO_2. Environmental Protection Agency. Accessed 11-04-2021 (public domain).

Air Pollution

7.2.2.1: Ozone Depletion

7.2.2.2: Acid Deposition

This page titled 7.2.2: Air Pollution is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

7.2.2.1: Ozone Depletion

The ozone depletion process begins when **CFCs** (chlorofluorocarbons) and other ozone-depleting substances (ODS) are emitted into the atmosphere (figure 7.2.2.1.*a*). CFC molecules are extremely stable, and they do not dissolve in rain. After a period of several years, ODS molecules reach the stratosphere, about 10 kilometers above the Earth's surface (figure 7.2.2.1.*b*). CFCs were used by industry as refrigerants, degreasing solvents, and propellants.

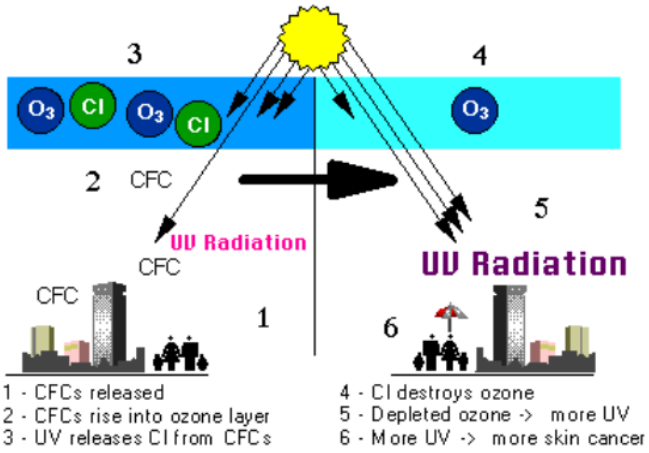


Figure 7.2.2.1.*a*: Strong ultraviolet (UV) light breaks apart the ozone-depleting substances (ODS). The process is as follows: (1) chlorofluorocarbons (CFCs) released, (2) CFCs rise into ozone layer in the stratosphere, which contains ozone (O_3), (3) UV radiation releases chlorine (Cl) from CFCs, (4) Cl destroys ozone, (5) depleted ozone allows more UV radiation to pass through the atmosphere, and (6) more UV radiation causes more skin cancer. Some ODS, including CFCs, hydrochlorofluorocarbons (HCFCs), carbon tetrachloride, and methyl chloroform release Cl atoms. Other ODS, including halons and methyl bromide, release bromine (Br) atoms. It is these atoms that actually destroy ozone, not the intact ODS molecule. It is estimated that one chlorine atom can destroy over 100,000 ozone molecules before it is removed from the stratosphere. Credit: NASA GSFC.

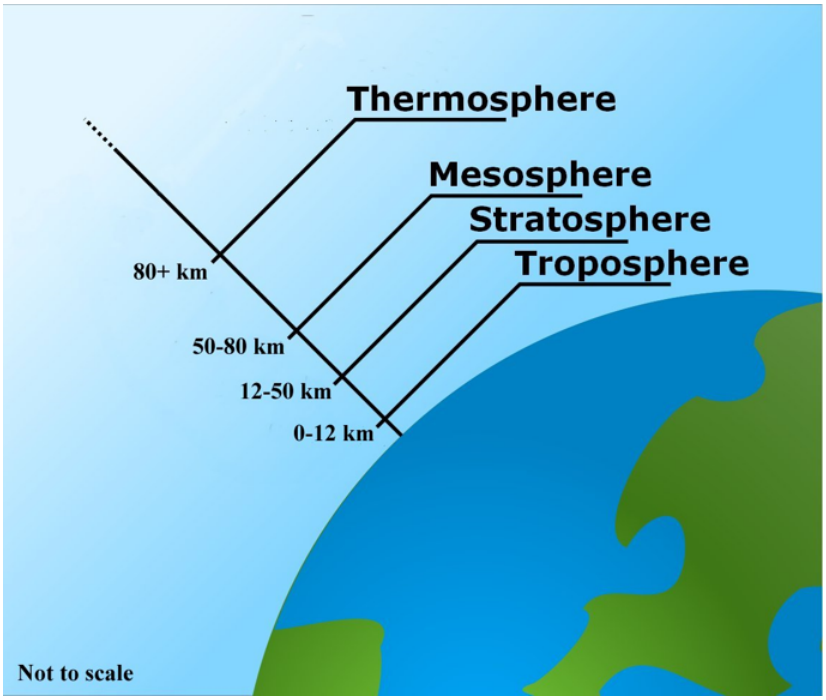


Figure 7.2.2.1.*b*: The layers of the atmosphere. the troposphere is the closest to the Earth's surface (0-12 km). Next, is the stratosphere (12-50 km), mesophere (50-80 km), and thermosphere (80+ km). The outermost layer (the exosphere) is not shown. **Ground-level ozone** in the troposphere is a form of air pollution, but the **ozone layer** in the stratosphere helps filter UV rays. Image by GFDL (CC-BY-SA).

Ozone (O_3) is constantly produced and destroyed in a natural cycle, as shown in figure 7.2.2.1.*c*. However, the overall amount of ozone is essentially stable. This balance can be thought of as a stream's depth at a particular location. Although individual water

molecules are moving past the observer, the total depth remains constant. Similarly, while ozone production and destruction are balanced, ozone levels remain stable. This was the situation until the past several decades. Large increases in stratospheric ODS, however, have upset that balance. In effect, they are removing ozone faster than natural ozone creation reactions can keep up. Therefore, ozone levels fall.

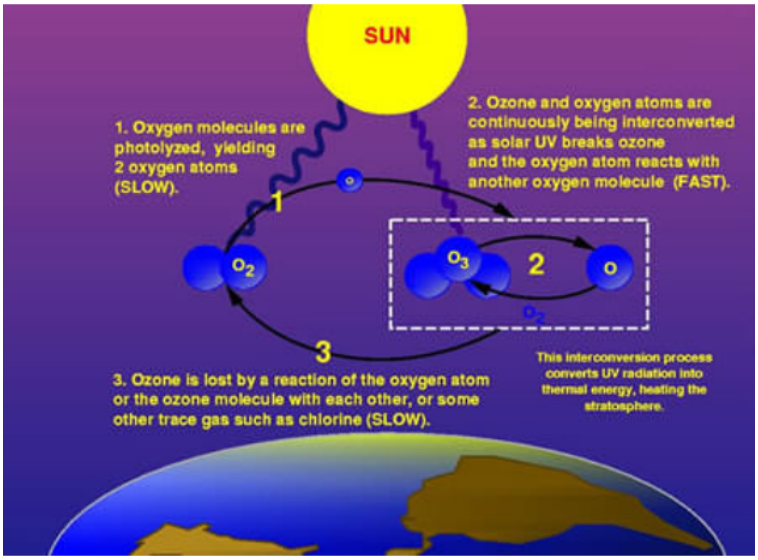


Figure 7.2.2.1.*c*: Because ozone (O_3) filters out a harmful type of ultraviolet radiation (UVB), less ozone means higher UVB levels at the surface. The more the depletion, the larger the increase in incoming UVB. UVB has been linked to skin cancer, cataracts, damage to materials like plastics, and harm to certain crops and marine organisms. Although some UVB reaches the surface even without ozone depletion, its harmful effects will increase as a result of this problem. The process of the ozone layer filtering UV radiation is as follows: (1) Gaseous oxygen molecules (O_2) are photolyzed (split), yielding two oxygen atoms. This is a slow process. (2) Ozone and oxygen atoms are continuously being interconverted as solar UV breaks O_3 into O_2 and a single oxygen atom (O). The oxygen atom (O) reacts with another O_2 molecule to form O_3. This interconversion is a fast process and converts UV radiation into thermal energy, heating the stratosphere. (3) Ozone is lost by a reaction of the oxygen atom or the ozone molecule with each other, or some other trace gas such as chlorine. This is a slow process.

Policies to Reduce Ozone Destruction

One success story in reducing pollutants that harm the atmosphere concerns ozone-destroying chemicals. In 1973, scientists calculated that CFCs could reach the stratosphere and break apart. This would release chlorine atoms, which would then destroy ozone. Based only on their calculations, the United States and most Scandinavian countries banned CFCs in spray cans in 1978. More confirmation that CFCs break down ozone was needed before more was done to reduce production of ozone-destroying chemicals. In 1985, members of the British Antarctic Survey reported that a 50% reduction in the ozone layer had been found over Antarctica in the previous three springs.

Two years after the British Antarctic Survey report, the “Montreal Protocol on Substances that Deplete the Ozone Layer” was ratified by nations all over the world. The **Montreal Protocol** controls the production and consumption of 96 chemicals that damage the ozone layer. CFCs have been mostly phased out since 1995, although they were used in developing nations until 2010. Some of the less hazardous substances will not be phased out until 2030. The Protocol also requires that wealthier nations donate money to develop technologies that will replace these chemicals.

Because CFCs take many years to reach the stratosphere and can survive there a long time before they break down, the ozone hole did not immediately disappear after CFC emissions were reduced; however, it has been shrinking (figure 7.2.2.1.*d*).

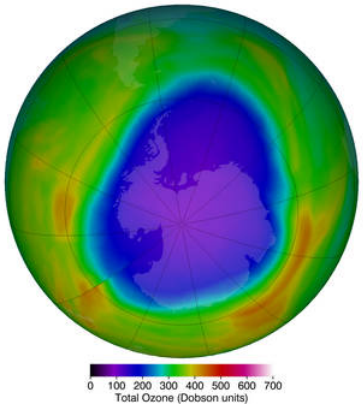


Figure 7.2.2.1.*d*: The Antarctic ozone hole that occurs annually in September and October during the Southern Hemisphere spring typically sees much lower ozone levels in than the Arctic. The purples and deep blues show the extent of low ozone levels on October 12, 2018, when they dropped to 104 Dobson units. Image and caption by NASA's Goddard Space Flight Center (public domain).

Health and Environmental Effects of Ozone Layer Depletion

There are three types of UV light: UVA, UVB, and UVC. Reductions in stratospheric ozone levels will lead to higher levels of UVB reaching the Earth's surface. The sun's output of UVB does not change; rather, less ozone means less protection, and hence more UVB reaches the Earth. Studies have shown that in the Antarctic, the amount of UVB measured at the surface can double during the annual ozone hole.

Laboratory and epidemiological studies demonstrate that UVB causes non-melanoma skin cancers and plays a major role in malignant melanoma development. In addition, UVB has been linked to cataracts, a clouding of the eye's lens. All sunlight contains some UVB, even with normal stratospheric ozone levels. Therefore, it is always important to protect your skin and eyes from the sun. Ozone layer depletion increases the amount of UVB and the risk of health effects.

UVB is generally harmful to cells, and therefore all organisms. UVB cannot penetrate into an organism very far and thus tends to only impact skin cells. Microbes like bacteria, however, are composed of only one cell and can therefore be harmed by UVB,

Attribution

Modified by Melissa Ha from Ozone Depletion from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 7.2.2.1: Ozone Depletion is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **20.2.1: Ozone Depletion** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

7.2.2.2: Acid Deposition

Acid deposition is a term referring to a mixture of wet and dry deposition (deposited material) from the atmosphere containing higher than normal amounts of nitric and sulfuric acids. The precursors, or chemical forerunners, of acid deposition formation result from both natural sources, such as volcanoes and decaying vegetation, and man-made sources, primarily emissions of sulfur dioxide (SO_2) and nitrogen oxides (NO_x) resulting from fossil fuel combustion (figure 7.2.2.2.a). Acid deposition occurs when these gases react in the atmosphere with water, oxygen, and other chemicals to form various acidic compounds. The result is a mild solution of sulfuric acid and nitric acid. When sulfur dioxide and nitrogen oxides are released from power plants and other sources, prevailing winds blow these compounds across state and national borders, sometimes over hundreds of miles.

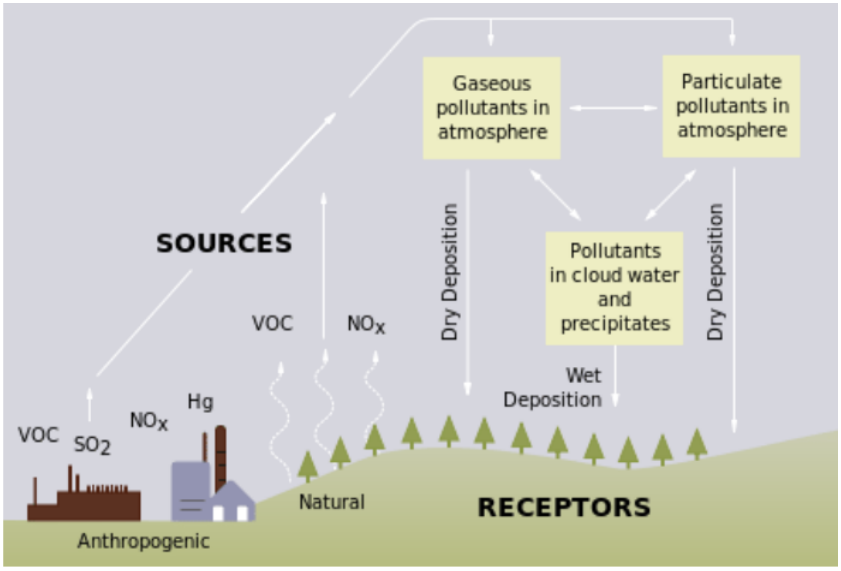
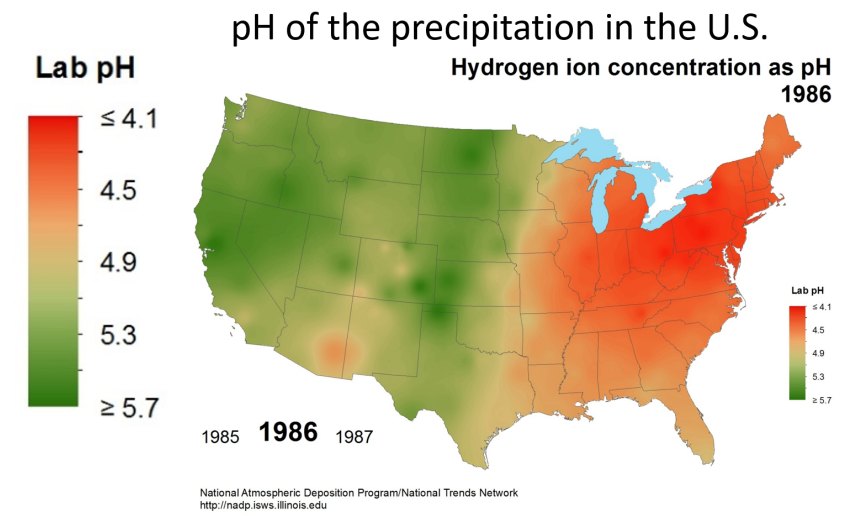


Figure 7.2.2.2.a: This diagram shows both anthropogenic and natural sources of acid deposition. Volatile organic compounds (VOCs), sulfur dioxide (SO_2), nitrogen oxides (NO_X), and mercury (Hg) are released, but only SO_2 and NO_X react to form acids. Three categories of pollutants interact with each other. These are gaseous pollutants in the atmosphere, particulate pollutants in the atmosphere, and pollutants in cloud water and precipitates. The former two categories form dry deposition, and the latter forms wet deposition. Acid deposition (both dry and wet deposition) reaches natural areas, which could be terrestrial (as pictured) or aquatic.

Acid deposition is measured using a scale called “pH.” The lower a substance’s pH, the more acidic it is. Pure water has a pH of 7.0. However, normal rain is slightly acidic because carbon dioxide (CO_2) dissolves into it forming weak carbonic acid, giving the resulting mixture a pH of approximately 5.6 at typical atmospheric concentrations of CO_2. The average pH of precipitation in the United States in 1980 was 4.6, but it has been gradually increasing due to reductions in sulfur dioxide and nitrogen oxide emissions (figure 7.2.2.2.b).



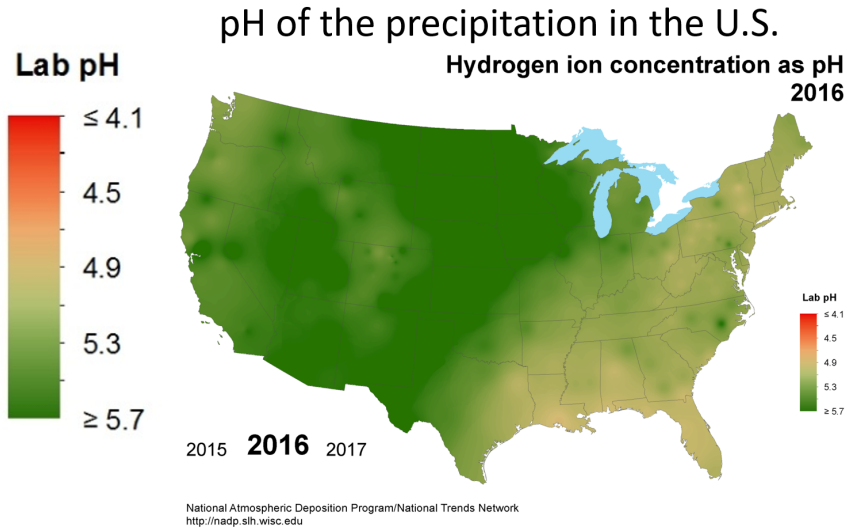


Figure 7.2.2.2.*b*: pH measures how acidic or basic a solution is, which depends on the concentration of hydrogen ions. Precipitation is normally slightly acidic, with a pH of around 5.6. (Anything below pH = 7 is considered acidic.) pH of precipitation was even more acidic in much of the East Coast of the U.S. in 1986 (acid deposition) with parts of Michigan, Ohio, Pennsylvania, and New York state reaching a pH of less than or equal to 4.1. The pH of the western United States in 1986 was closer to 5.0. Due to efforts to reduce air pollution, the pH of precipitation has moved closer to normal. By 2016, the pH of the eastern United States had increased to around 4.9, and the pH of the western United States had mostly returned to normal, with many regions having a pH of greater than or equal to 5.7. Images by National Atmospheric Deposition Program (public domain).

Effects of Acid Deposition

Acid deposition causes **acidification** of lakes and streams and contributes to the damage of trees at high elevations (for example, red spruce trees above 2,000 feet) and many sensitive forest soils. In addition, acid deposition accelerates the decay of building materials and paints, including irreplaceable buildings, statues, and sculptures that are part of our nation's cultural heritage. Prior to falling to the earth, sulfur dioxide (SO_2) and nitrogen oxide (NO_x) gases and their particulate matter derivatives—sulfates and nitrates—contribute to visibility degradation and harm public health.

The **ecological** effects of acid deposition are most clearly seen in the aquatic, or water, environments, such as streams, lakes, and marshes. Most lakes and streams have a pH between 6 and 8, although some lakes are naturally acidic even without the effects of

acid deposition. Acid deposition primarily affects sensitive bodies of water, which are located in watersheds whose soils have a limited ability to neutralize acidic compounds (called “buffering capacity”). Lakes and streams become acidic (i.e., the pH value goes down) when the water itself and its surrounding soil cannot buffer the acid deposition enough to neutralize it. In areas where buffering capacity is low, acid deposition releases aluminum from soils into lakes and streams; aluminum is highly toxic to many species of aquatic organisms. Acid deposition causes slower growth, injury, or death of **forests**. Of course, acid deposition is not the only cause of such conditions. Other factors contribute to the overall stress of these areas, including air pollutants, insects, disease, drought, or very cold weather. In most cases, in fact, the impacts of acid deposition on trees are due to the combined effects of acid deposition and these other environmental stressors.

Acid deposition contributes to the corrosion of **metals** (such as bronze) and the deterioration of paint and stone (such as marble and limestone). These effects significantly reduce the societal value of buildings, bridges, cultural objects (such as statues, monuments, and tombstones), and cars (figure 7.2.2.2.*c*).



Figure 7.2.2.2.*c*: A gargoyle that has been damaged by acid deposition.

Sulfates and nitrates that form in the atmosphere from sulfur dioxide (SO_2) and nitrogen oxides (NO_x) emissions contribute to **visibility impairment,** meaning we cannot see as far or as clearly through the air. The pollutants that cause acid deposition—sulfur dioxide (SO_2) and nitrogen oxides (NO_x)—damage **human health**. These gases interact in the atmosphere to form fine sulfate and nitrate particles that can be transported long distances by winds and inhaled deep into people’s lungs. Fine particles can also penetrate indoors. Many scientific studies have identified a relationship between elevated levels of fine particles and increased illness and premature death from heart and lung disorders, such as asthma and bronchitis.

Attribution

Modified by Melissa Ha from Acid Rain from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 7.2.2.2: Acid Deposition is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **20.2.2: Acid Deposition** by Melissa Ha and Rachel Schleiger is licensed CC BY 4.0.

7.2.3: Data Dive- Plastic in World Oceans

Overview

Our World in Data (OWID) is a scientific online publication that focuses on using research and data to help tackle global concerns such as poverty, disease, hunger, climate change, war, and inequitable treatment of our world's most vulnerable and unstable communities. Their website is particularly known for publishing a variety of graphs, some even interactive, to present the research that helps explain causes and consequences of global concerns to the public. One example graph, seen below, illustrates the pathway by which plastic pollution enters the world's oceans:

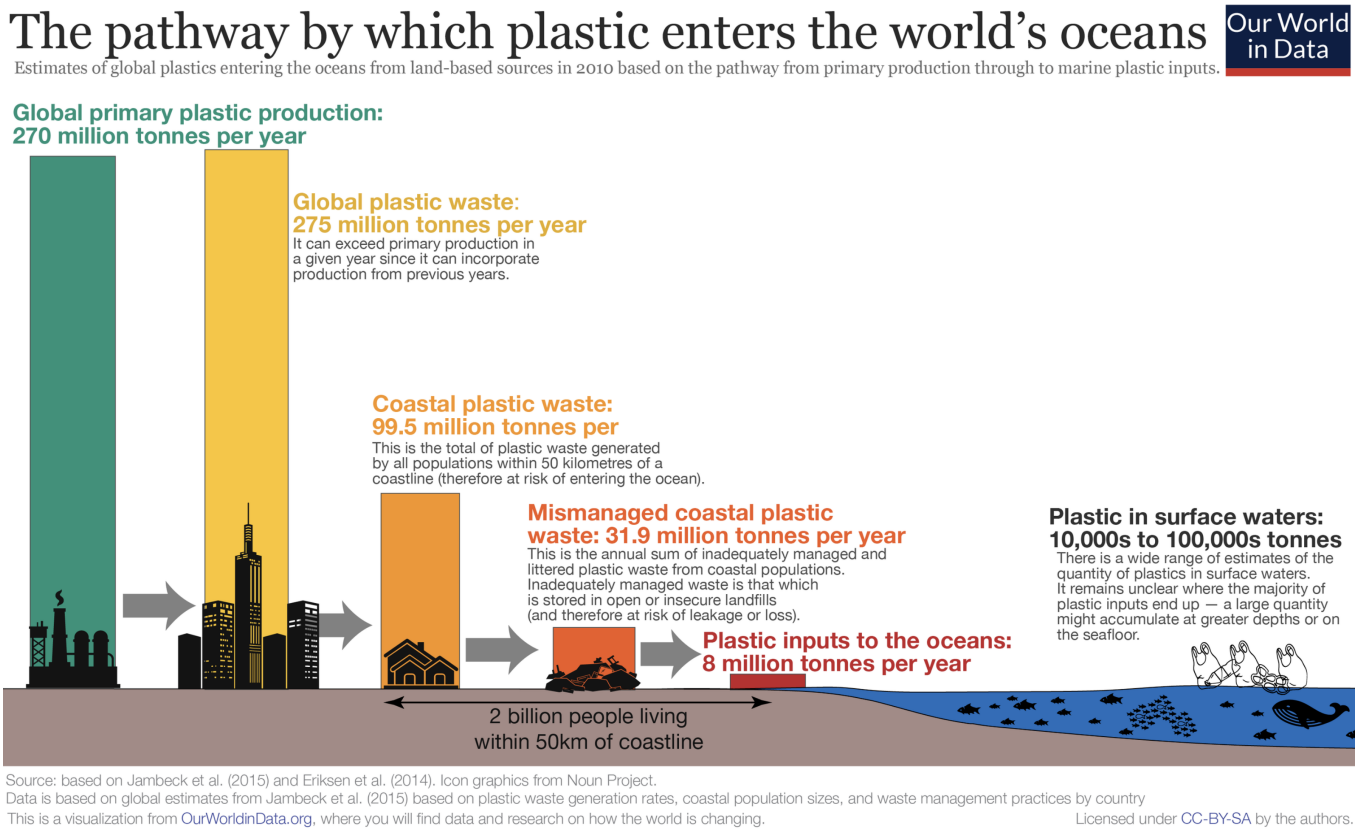


Figure 7.2.3.*a*: Plastic pathway into the ocean by million tonnes per year. Graph by Our World in Data (CC-BY-SA).

Questions

1. Reflect on how this bar graph is different than a standard bar graph. What do you like or not like about it?
2. What is this graph is illustrating about plastic waste?
3. Note that the global plastic waste is higher than the global plastic production. What does that tell you about the amount of plastic not yet classified as "waste"?
4. Provide a brief reflection for how these results make you feel and what they make you think about?
5. How can the results of this graph be used to inform future policies about plastic waste?

Raw Data From Above Graph(s)

Table 7.2.3.*a***:** Raw data for plastic pathway into the ocean by million tonnes per year. Graph by Our World in Data (CC-BY-SA).

Waste Category	Tonnes Per Year
Global primary waste production	270 million
Global plastic waste	275 million
Coastal plastic waste	99.5 million
Mismanaged coastal plastic waste	31.9 million
Plastic inputs to the oceans	8 million
Plastic in surface waters	10,000s to 100,000s

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 7.2.3: Data Dive- Plastic in World Oceans is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **20.3: Data Dive- Plastic in World Oceans** by Melissa Ha and Rachel Schleiger has no license indicated.

7.2.4: Review

Summary

After completing this chapter you should be able to...

- Distinguish between point source and nonpoint source water pollution.
- Name and describe common water pollutants, including chemical, biological, and physical pollutants.
- Explain the mechanism of eutrophication.
- Summarize the current state of wastewater treatment globally.
- Describe the process of wastewater treatment, including pretreatment, primary treatment, secondary treatment, tertiary treatment, and disinfection and discharge.
- Summarize strategies for reducing water pollution.
- Identify sources of air pollution.
- List common air pollutants.
- Explain how CFCs caused ozone depletion, and global efforts to address this issue.
- Describe the causes and consequences of acid deposition.

Water pollution may arise from a single origin (**point source pollution**), or it may arise from multiple dispersed sources throughout the **watershed** (**nonpoint source pollution**). Water pollutants may be **chemical**, **biological**, or **physical**. **Oxygen-demanding waste** increases **biological oxygen demand** and causes **hypoxia**, depriving aquatic organisms of oxygen. This results from **eutrophication**, in which excess nutrients cause **algal blooms**.

Pathogens are the most deadly form of water pollution. They cause waterborne diseases, killing 485,000 people every year. Resolution of the global water pollution crisis requires multiple approaches to improve the quality of fresh water. The best strategy for addressing this problem is proper **wastewater treatment**. Strategies to reduce water pollution in general include the **Clean Water Act**, **remediation**, and **watershed management**.

Air pollution can be thought of as gaseous and particulate contaminants that are present in the Earth's atmosphere. Chemicals discharged into the air that have a direct impact on the environment are called **primary pollutants**. These primary pollutants sometimes react with other chemicals in the air to produce **secondary pollutants**. The commonly found air pollutants, known as **criteria air pollutants**, are particle pollution, ground-level ozone, carbon monoxide, sulfur oxides, nitrogen oxides, and lead. These pollutants can harm health and the environment and cause property damage.

The ozone depletion process begins when **chlorofluorocarbons** (CFCs) and other **ozone-depleting substances** (ODS) are emitted into the atmosphere. Reductions in stratospheric ozone levels lead to higher levels of harmful ultraviolet radiation, particularly UVB, reaching the Earth's surface. The sun's output of UVB does not change; rather, less ozone means less protection, and hence more UVB reaches the Earth. The **Montreal Protocol** is an international effort to phase out CFCs, and has been successful in limiting ozone depletion.

Acid deposition occurs when certain air pollutants react with atmosphere to produce nitric and sulfuric acids. It can reach Earth as various forms of precipitation or as dry particles that later react to form acid. The precursors of acid deposition result from both natural sources, such as volcanoes and decaying vegetation, and **anthropogenic** (human) sources, primarily emissions of sulfur dioxide (SO_2) and nitrogen oxides (NOx) resulting from fossil fuel combustion. Acid deposition causes acidification of lakes and streams, contributes to the damage of trees and many sensitive forest soils. In addition, acid deposition accelerates the decay of building materials and paints, contributes to the corrosion of metals and damages human health. However, the severity of acid deposition has declined due ot regulations and technologies that limit air pollution.

Attribution

Modified by Melissa Ha from Water Availability and Use and Air Pollution, Climate Change, & Ozone Depletion from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 7.2.4: Review is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **20.4: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY 4.0.

SECTION OVERVIEW

7.3: Climate Change

Chapter Hook

The iconic Grand Canal of Venice, Italy, doesn't need high quality instruments to note effects of climate change. The Grand Canal was built in a way that the water levels are directly correlated with the sea level. As such, the steps leading into the canal have become rulers for sea level changes over the hundreds of years of their existence. The last step leading into the canal directly reflected the sea level at the time of creation and currently, the water level is 3 feet above this. In the last quarter century, sea levels have been rising much faster than previously measured. This puts Venice in a difficult position moving into a future where scientists predict the sea level to continue to rise, estimating that the city will be underwater by 2100, along with many other coastal cities. Unfortunately, rising seas are only one of the many predicted outcomes of global climate change.



Figure 7.3.*a*: Flooding in Venice, Italy. Image by WorldIslandInfo.com in Wikimedia Commons (CC-BY2.0).

Climate change refers to any significant change in the measures of climate lasting for an extended period of time. In other words, climate change includes major changes in temperature, precipitation, or wind patterns, among other effects, that occur over several decades or longer. **Global warming** refers to the recent and ongoing rise in global average temperature near Earth's surface. It is caused mostly by increasing concentrations of greenhouse gases in the atmosphere. Global warming is causing climate patterns to change. However, global warming itself represents only one aspect of climate change.

A common misconception about global climate change is that a specific weather event occurring in a particular region (for example, a very cool week in June in central Indiana) is evidence of global climate change. However, a cold week in June is a weather-related event and not a climate-related one. These misconceptions often arise because of confusion over the terms climate and weather.

Climate refers to the long-term, predictable atmospheric conditions of a specific area. The climate of a biome is characterized by having consistent temperature and annual rainfall ranges. Climate does not address the amount of rain that fell on one particular day in a biome or the colder-than-average temperatures that occurred on one day. In contrast, weather refers to the conditions of the

atmosphere during a short period of time. Weather forecasts are usually made for 48-hour cycles. Long-range weather forecasts are available but can be unreliable.

To better understand the difference between climate and weather, imagine that you are planning an outdoor event in northern Wisconsin. You would be thinking about *climate* when you plan the event in the summer rather than the winter because you have long-term knowledge that any given Saturday in the months of May to August would be a better choice for an outdoor event in Wisconsin than any given Saturday in January. However, you cannot determine the specific day that the event should be held on because it is difficult to accurately predict the weather on a specific day. Climate can be considered “average” weather.

7.3.1: The Greenhouse Effect and Climate Change

7.3.2: Data Dive- Flooding in Venice Italy

7.3.3: Review

Attributions

Modified by Melissa Ha and Rachel Schleiger from the following sources:

- Climate and the Effects of Global Climate Change from *General Biology* by OpenStax (licensed under CC-BY)
- Climate Change from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 7.3: Climate Change is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

7.3.1: The Greenhouse Effect and Climate Change

Earth's Temperature is a Balancing Act

Earth's temperature depends on the balance between energy entering and leaving the planet. When incoming energy from the sun is absorbed, Earth warms. When the sun's energy is reflected back into space, Earth avoids warming. When energy is released from Earth into space, the planet cools. Many factors, both natural and human, can cause changes in Earth's energy balance, including:

- Changes in the greenhouse effect, which affects the amount of heat retained by Earth's atmosphere;
- Variations in the sun's energy reaching Earth;
- Changes in the reflectivity of Earth's atmosphere and surface.

Scientists have pieced together a picture of Earth's climate, dating back hundreds of thousands of years, by analyzing a number of indirect measures of climate such as ice cores, tree rings, glacier size, pollen counts, and ocean sediments. Scientists have also studied changes in Earth's orbit around the sun and the activity of the sun itself.

The historical record shows that the climate varies naturally over a wide range of time scales. In general, climate changes prior to the Industrial Revolution in the 1700s can be explained by natural causes, such as changes in solar energy, volcanic eruptions, and natural changes in greenhouse gas (GHG) concentrations. Recent changes in climate**,** however, cannot be explained by natural causes alone. Research indicates that natural causes are very unlikely to explain most observed warming, especially warming since the mid-20th century. Rather, human activities, especially our combustion of fossil fuels, explains the current warming (figure 7.3.1.*a*). The scientific consensus is clear: through alterations of the carbon cycle, humans are changing the global climate by increasing the effects of something known as the greenhouse effect.

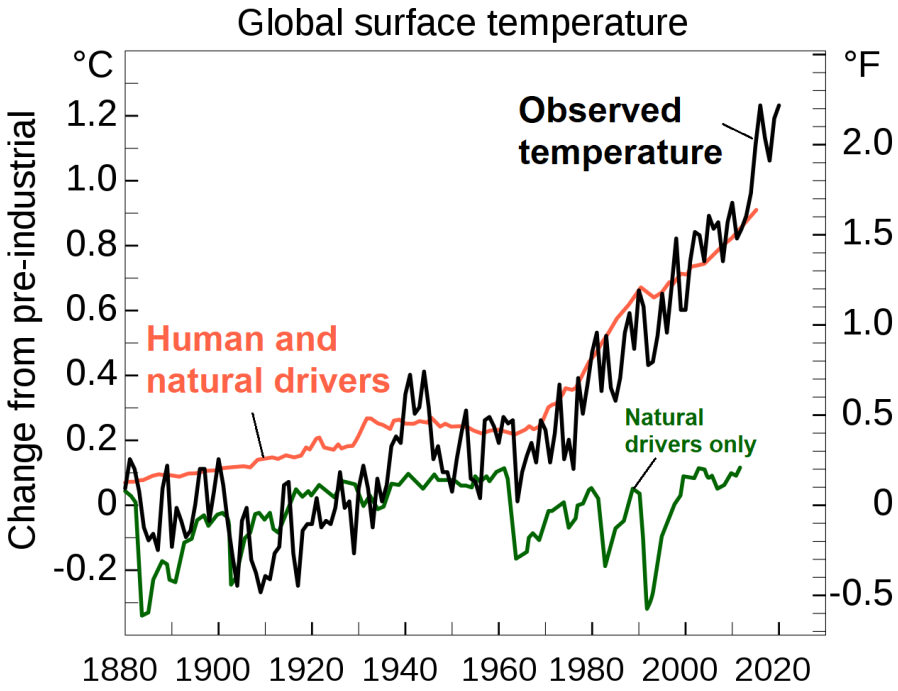


Figure 7.3.1.*a*: This graph shows the predicted temperatures from two climate models and observed temperatures from 1880 to 2020. The first model considered only natural factors that could influence temperature and is represented by the green line (bottom). It shows some fluctuations in temperature by no overall increase or decrease. The second model considered both human and natural factors and is represented by the orange line (top). It shows an overall increase in temperature. Actual observations (black, jagged line; middle) more closely match the second model. Overall, temperature has increased about 1.2 degrees Celsius (2.1 degrees Fahrenheit) since pre-industrial times. Image relabeled from Efbrazil (CC-BY-SA).

The Greenhouse Effect Causes the Atmosphere to Retain Heat

Gardeners that live in moderate or cool environments use greenhouses because they trap heat and create an environment that is warmer than outside temperatures. This is great for plants that like heat, or are sensitive to cold temperatures, such as tomato and

pepper plants. Greenhouses contain glass or plastic that allow visible light from the sun to pass. This light, which is a form of energy, is absorbed by plants, soil, and surfaces and heats them. Some of that heat energy is then radiated outwards in the form of infrared radiation, a different form of energy. Unlike with visible light, the glass of the greenhouse blocks the infrared radiation, thereby trapping the heat energy, causing the temperature within the greenhouse to increase.

The same phenomenon happens inside a car on a sunny day. Have you ever noticed how much hotter a car can get compared to the outside temperature? Light energy from the sun passes through the windows and is absorbed by the surfaces in the car such as seats and the dashboard. Those warm surfaces then radiate infrared radiation, which cannot pass through the glass. This trapped infrared energy causes the air temperatures in the car to increase. This process is commonly known as the **greenhouse effect**.

The video below made for kids, but provides a clear and simple introduction to the greenhouse effect.



The greenhouse effect also happens with the entire Earth. Of course, our planet is not surrounded by glass windows. Instead, the Earth is wrapped with an atmosphere that contains **greenhouse gases** (GHGs). Much like the glass in a greenhouse, GHGs allow incoming visible light energy from the sun to pass, but they block infrared radiation that is radiated from the Earth towards space (figure 7.3.1.*b*). In this way, they help trap heat energy that subsequently raises air temperature. Being a greenhouse gas is a physical property of certain types of gases; because of their molecular structure they absorb wavelengths of infrared radiation, but are transparent to visible light. Some notable greenhouse gases are water vapor (H_2O), carbon dioxide (CO_2), and methane (CH_4). GHGs act like a blanket, making Earth significantly warmer than it would otherwise be. Scientists estimate that average temperature on Earth would be -18° C without naturally-occurring GHGs.

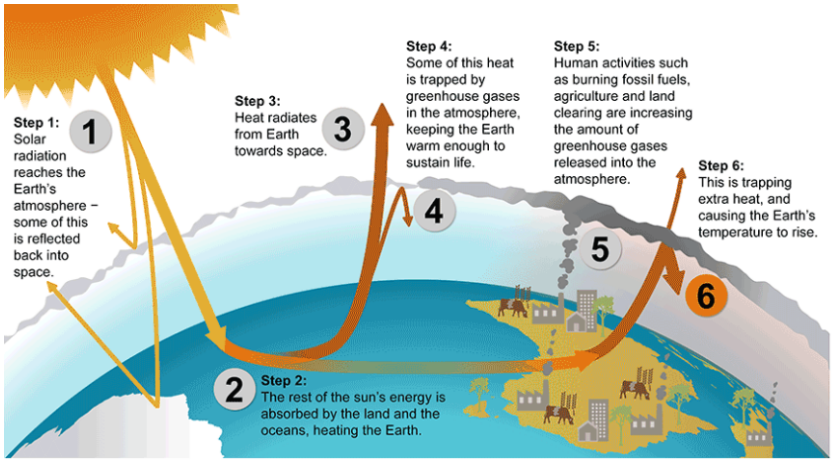


Figure 7.3.1.*b*: The enhanced greenhouse effect. Step 1: Some solar radiation reaches the Earth's surface, and some is reflected back into space. Step 2: The rest of the sun's energy is absorbed by the land and the oceans, heating the Earth. Step 3: Heat radiates from Earth towards space. Step 4: Some of the heat is trapped by greenhouse gases in the atmosphere, warming the Earth. Step 5: Human activities such as burning fossil fuels, agriculture, and land clearing have increased the concentrations of greenhouse gases in the atmosphere. Step 6: This is trapping extra heat, causing the Earth's temperature to rise.

What is Global Warming?

Global warming refers to the recent and ongoing rise in global average temperature near Earth's surface. It is caused mostly by increasing concentrations of greenhouse gases in the atmosphere. Global warming is causing climate patterns to change. However, global warming itself represents only one aspect of climate change.

What is Climate Change?

Climate change refers to any significant change in the measures of climate lasting for an extended period of time. In other words, climate change includes major changes in temperature, precipitation, or wind patterns, among other effects, that occur over several decades or longer.

The Main Greenhouse Gasses

The most important GHGs directly emitted by humans include CO_2 and methane. **Carbon dioxide** (CO_2) is the primary greenhouse gas that is contributing to recent global climate change. CO_2 is a natural component of the carbon cycle, involved in such activities as photosynthesis, respiration, volcanic eruptions, and ocean-atmosphere exchange. Human activities, primarily the burning of fossil fuels and changes in land use, release very large amounts of CO_2 to the atmosphere, causing its concentration in the atmosphere to rise.

Atmospheric CO_2 concentrations have increased by 45% since pre-industrial times, from approximately 280 parts per million (ppm) in the 18th century to 409.8 ppm in 2019 (figure 7.3.1.c). The current CO_2 level is higher than it has been in at least 800,000 years, based on evidence from ice cores that preserve ancient atmospheric gases (figure 7.3.1.$d-f$). Human activities currently release over 30 billion tons of CO_2 into the atmosphere every year. While some volcanic eruptions released large quantities of CO_2 in the distant past, the U.S. Geological Survey (USGS) reports that human activities now emit more than 135 times as much CO_2 as volcanoes each year. This human-caused build-up of CO_2 in the atmosphere is like a tub filling with water, where more water flows from the faucet than the drain can take away.

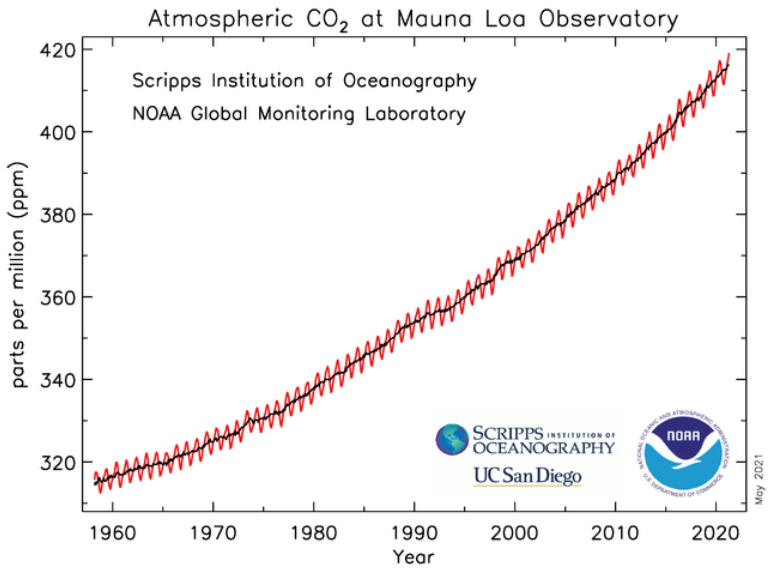


Figure 7.3.1.c: Monthly mean carbon dioxide concentration in parts per million (ppm) measured at Mauna Loa Observatory, Hawaii. The carbon dioxide data on Mauna Loa constitute the longest record of direct measurements of CO_2 in the atmosphere. The fluctuating red line represents the monthly mean values, centered on the middle of each month. Carbon dioxide concentrations dip every summer due to increased photosynthesis. The smoother black line represents the same, after correction for the average seasonal cycle. Image and caption (modified) by NOAA (public domain).

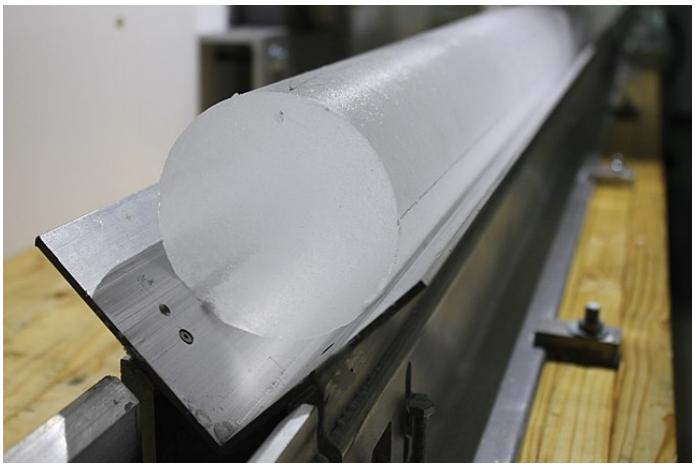


Figure 7.3.1.*d*: The EastGRIP ice core, which was drilled from the Greenland Ice Sheet. You can read more here. Image by Helle Astrid Kjær (CC-BY).

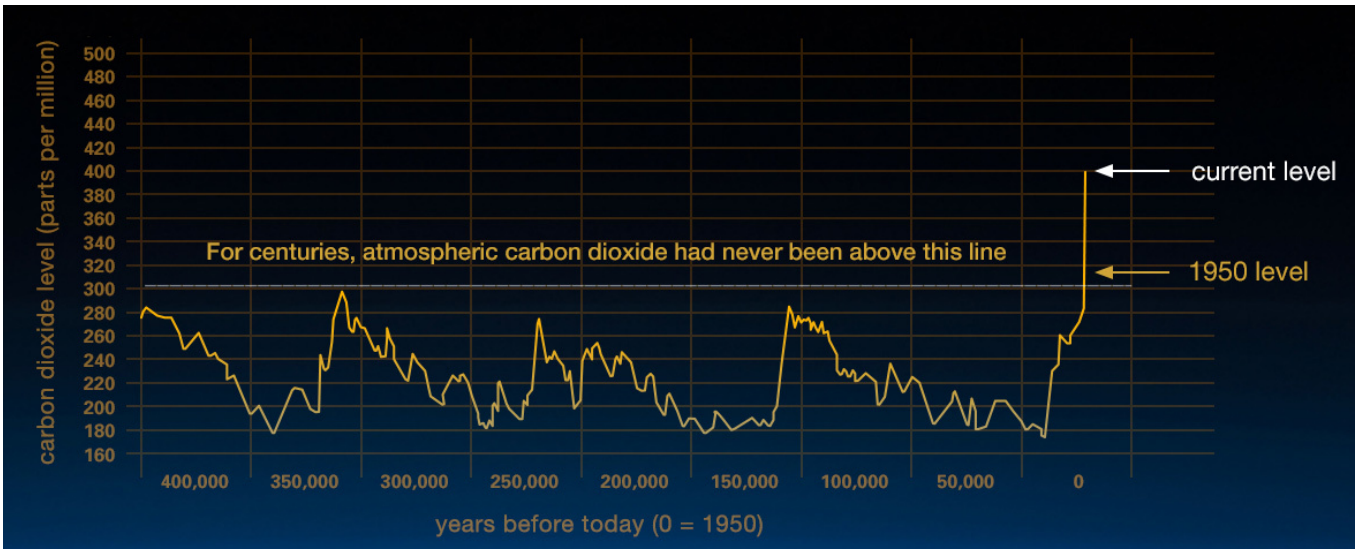


Figure 7.3.1.*e*: This graph, based on the comparison of atmospheric samples contained in ice cores and more recent direct measurements, provides evidence that atmospheric CO2 has increased since the Industrial Revolution. On the x-axis are the years before today (0 = 1950). It begins 400,000 years prior to 1950. On the y-axis is the carbon dioxide level in parts per million. Carbon dioxide levels have fluctuated over the years, but they never exceeded 300 parts per million until 1950. In 2018, carbon dioxide levels reached 409.8 ppm. (This graph is a few years old and shows the current CO_2 level at 400 ppm; Credit: Vostok ice core data/J.R. Petit et al.; NOAA Mauna Loa CO2 record.)

https://bio.libretexts.org/@go/page/109240

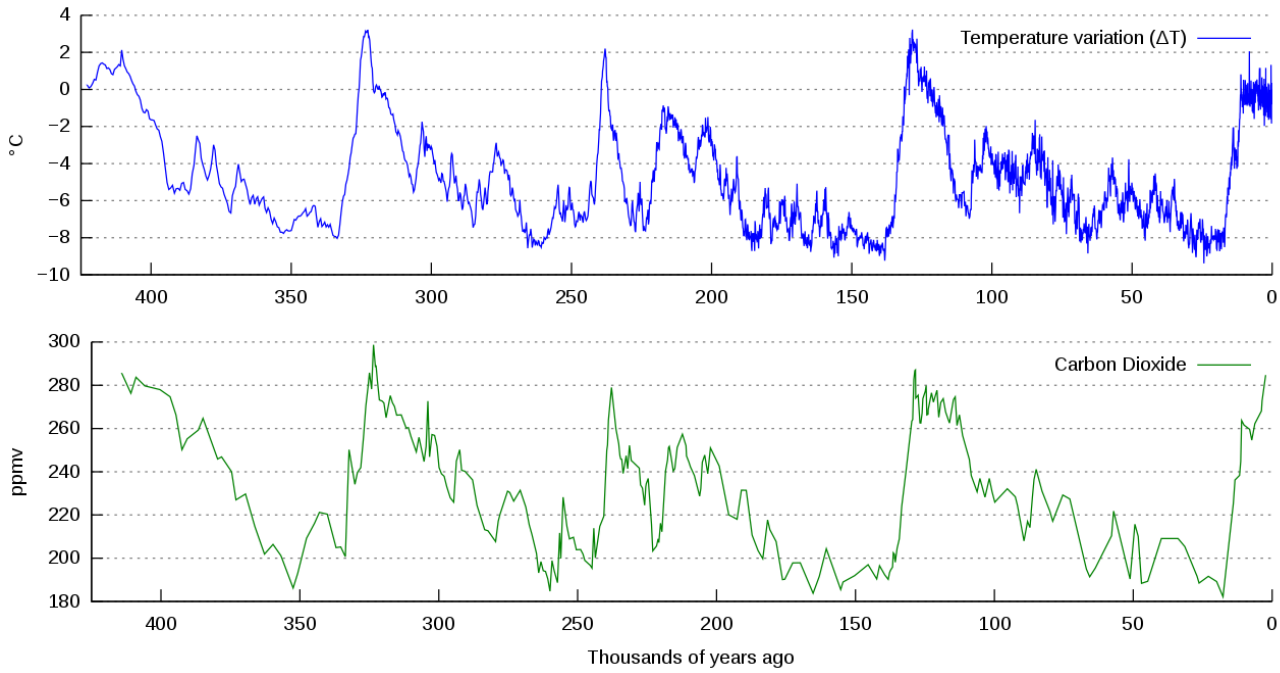


Figure 7.3.1.*f*: Graph of change in temperature in degrees Celsius (top, blue line) and carbon dioxide concentration in parts per million by volume (bottom, green line) measured from the Vostok, Antarctica ice core. These have been associated for over 400,000 years. As carbon dioxide concentration increased, so did temperature. As carbon dioxide concentration decreased, so did temperature. These data were collected in 1999. Since then, carbon dioxide concentrations have increased to 409.8 ppm (2019 average). Image and caption (modified) by NOAA/Autopilot (CC-BY-SA)

Other Greenhouse Gasses

Although this concentration is far less than that of CO_2, **methane** (CH_4) is 28 times as potent a greenhouse gas. Methane is produced when bacteria break down organic matter under anaerobic conditions and can be released due to natural or anthropogenic processes. Anaerobic conditions can happen when organic matter is trapped underwater (such as in rice paddies) or in the intestines of herbivores. Anthropogenic causes now account for 60% of total methane release. Examples include agriculture, fossil fuel extraction and transport, mining, landfill use, and burning of forests. Specifically, raising cattle releases methane due to fermentation in their rumens produces methane that is expelled from their GI tract. Methane is more abundant in Earth's atmosphere now than at any time in at least the past 650,000 years, and CH_4 concentrations increased sharply during most of the 20th century. They are now more than two and-a-half times pre-industrial levels (1.9 ppm), but the rate of increase has slowed considerably in recent decades.

Water vapor is the most abundant greenhouse gas and also the most important in terms of its contribution to the natural greenhouse effect, despite having a short atmospheric lifetime. Some human activities can influence local water vapor levels. However, on a global scale, the concentration of water vapor is controlled by temperature, which influences overall rates of evaporation and precipitation. Therefore, the global concentration of water vapor is not substantially affected by direct human emissions.

Ground-level ozone (O_3), which also has a short atmospheric lifetime, is a potent greenhouse gas. Chemical reactions create ozone from emissions of nitrogen oxides and volatile organic compounds from automobiles, power plants, and other industrial and commercial sources in the presence of sunlight (as discussed in section 10.1). In addition to trapping heat, ozone is a pollutant that can cause respiratory health problems and damage crops and ecosystems.

Changes in the Sun's Energy Affect how Much Energy Reaches Earth

Climate can be influenced by natural changes that affect how much solar energy reaches Earth. These changes include changes within the sun and changes in Earth's orbit. Changes occurring in the sun itself can affect the intensity of the sunlight that reaches Earth's surface. The intensity of the sunlight can cause either warming (during periods of stronger solar intensity) or cooling (during periods of weaker solar intensity). The sun follows a natural 11-year cycle of small ups and downs in intensity, but the

effect on Earth's climate is small. Changes in the shape of Earth's orbit as well as the tilt and position of Earth's axis can also affect the amount of sunlight reaching Earth's surface.

Changes in the sun's intensity have influenced Earth's climate in the past. For example, the so-called "**Little Ice Age**" between the 17th and 19th centuries may have been partially caused by a low solar activity phase from 1645 to 1715, which coincided with cooler temperatures. The Little Ice Age refers to a slight cooling of North America, Europe, and probably other areas around the globe. Changes in Earth's orbit have had a big impact on climate over tens of thousands of years. These changes appear to be the primary cause of past cycles of ice ages, in which Earth has experienced long periods of cold temperatures (ice ages), as well as shorter interglacial periods (periods between ice ages) of relatively warmer temperatures.

Changes in solar energy continue to affect climate. However, solar activity has been relatively constant, aside from the 11-year cycle, since the mid-20th century and therefore does not explain the recent warming of Earth. Similarly, changes in the shape of Earth's orbit as well as the tilt and position of Earth's axis affect temperature on relatively long timescales (tens of thousands of years), and therefore cannot explain the recent warming.

Changes in Reflectivity Affect How Much Energy Enters Earth's System

When sunlight energy reaches Earth it can be reflected or absorbed. The amount that is reflected or absorbed depends on Earth's surface and atmosphere. Light-colored objects and surfaces, like snow and clouds, tend to reflect most sunlight, while darker objects and surfaces, like the ocean and forests, tend to absorb more sunlight. The term **albedo** refers to the amount of solar radiation reflected from an object or surface, often expressed as a percentage. Earth as a whole has an albedo of about 30%, meaning that 70% of the sunlight that reaches the planet is absorbed. Sunlight that is absorbed warms Earth's land, water, and atmosphere.

Albedo is also affected by aerosols. **Aerosols** are small particles or liquid droplets in the atmosphere that can absorb or reflect sunlight. Unlike greenhouse gases (GHGs), the climate effects of aerosols vary depending on what they are made of and where they are emitted. Those aerosols that reflect sunlight, such as particles from volcanic eruptions or sulfur emissions from burning coal, have a cooling effect. Those that absorb sunlight, such as black carbon (a part of soot), have a warming effect.

Natural changes in albedo, like the melting of sea ice or increases in cloud cover, have contributed to climate change in the past, often acting as feedbacks to other processes. Volcanoes have played a noticeable role in climate. Volcanic particles that reach the upper atmosphere can reflect enough sunlight back to space to cool the surface of the planet by a few tenths of a degree for several years. Volcanic particles from a single eruption do not produce long-term change because they remain in the atmosphere for a much shorter time than GHGs.

Human changes in land use and land cover have changed Earth's albedo. Processes such as deforestation, reforestation, desertification, and urbanization often contribute to changes in climate in the places they occur. These effects may be significant regionally, but are smaller when averaged over the entire globe.

Scientific Consensus: Global Climate Change is Real

The **Intergovernmental Panel on Climate Change** (IPCC) was created in 1988 by the United Nations Environment Programme and the World Meteorological Organization. It is charged with the task of evaluating and synthesizing the scientific evidence surrounding global climate change. The IPCC uses this information to evaluate current impacts and future risks, in addition to providing policymakers with assessments. These assessments are released about once every every six years. The most recent report, the 5th Assessment, was released in 2013. Hundreds of leading scientists from around the world are chosen to author these reports. Over the history of the IPCC, these scientists have reviewed thousands of peer-reviewed, publicly available studies. The scientific consensus is clear: global climate change is real and humans are very likely the cause for this change.

Additionally, the major scientific agencies of the United States, including the National Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA), also agree that climate change is occurring and that humans are driving it. In 2010, the US National Research Council concluded that "Climate change is occurring, is very likely caused by human activities, and poses significant risks for a broad range of human and natural systems". Many independent scientific organizations have released similar statements, both in the United States and abroad. This doesn't necessarily mean that every scientist sees eye to eye on each component of the climate change problem, but broad agreement exists that climate change is happening and is primarily caused by excess greenhouse gases from human activities. Critics of climate change, driven by ideology

instead of evidence, try to suggest to the public that there is no scientific consensus on global climate change. Such an assertion is patently false.

Current Status of Global Climate Change and Future Changes

Greenhouse gas concentrations in the atmosphere will continue to increase unless the billions of tons of anthropogenic emissions each year decrease substantially. Increased concentrations are expected to do the following:

- Increase Earth's average temperature (figure 7.3.1.*g*),
- Influence the patterns and amounts of precipitation,
- Reduce ice and snow cover, as well as permafrost,
- Raise sea level (figure 7.3.1.*h*),
- Increase the acidity of the oceans.

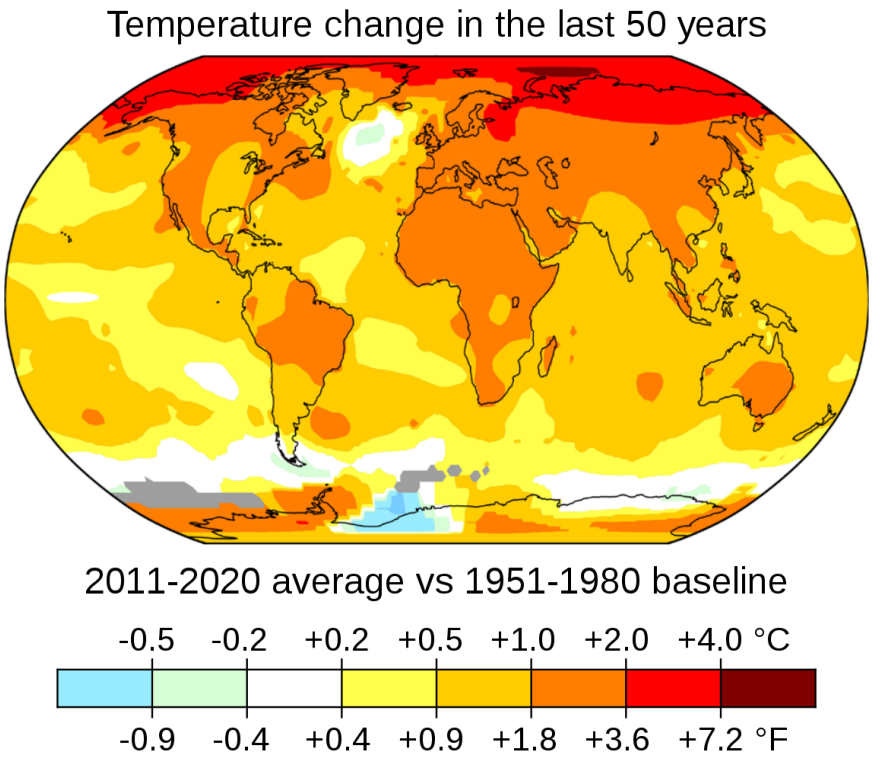


Figure 7.3.1.*g*: Temperature increases have been most pronounced in northern latitudes and over land masses. The colors represent the temperature difference between the 2011-2020 average and the 1951-1980 baseline, with warmer colors (yellow, orange, red) representing increases, and cool colors (green, blue) representing decreases. The image uses longer term averages of at least a decade to smooth out climate variability due to factors such as El Niño. Grey areas in the image have insufficient data for rendering. Image and caption (modified) from NASA's Scientific Visualization Studio/Eric Fisk (public domain).

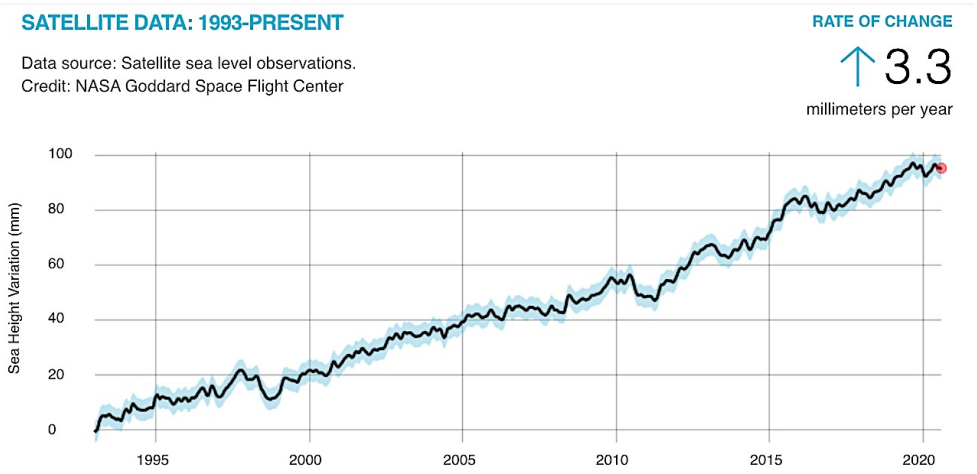


Figure 7.3.1.*h***:** Sea height variation (mm) over time. Sea height has increased about 3.3 millimeters per year on average since 1993. Data is from satellite sea level observations by the NASA Goddard Space Flight Center. Image by NASA (public domain).

These changes will impact our food supply, water resources, infrastructure, ecosystems, and even our own health. The magnitude and rate of future climate change will primarily depend on the following factors:

- The rate at which levels of greenhouse gas concentrations in our atmosphere continue to increase,
- How strongly features of the climate (e.g., temperature, precipitation, and sea level) respond to the expected increase in greenhouse gas concentrations,
- Natural influences on climate (e.g., from volcanic activity and changes in the sun's intensity) and natural processes within the climate system (e.g., changes in ocean circulation patterns).

Past and Present-day GHG Emissions Will Affect Climate Far into the Future

Many greenhouse gases stay in the atmosphere for long periods of time. As a result, even if emissions stopped increasing, atmospheric greenhouse gas concentrations would continue to remain elevated for hundreds of years. Moreover, if we stabilized concentrations and the composition of today's atmosphere remained steady (which would require a dramatic reduction in current greenhouse gas emissions), surface air temperatures would continue to warm. This is because the oceans, which store heat, take many decades to fully respond to higher greenhouse gas concentrations. The ocean's response to higher greenhouse gas concentrations and higher temperatures will continue to impact climate over the next several decades to hundreds of years.

Future Temperature Changes

Climate models project the following key temperature-related changes:

- Average global temperatures are expected to increase by 2°F to 11.5°F by 2100, depending on the level of future greenhouse gas emissions, and the outcomes from various climate models.
- By 2100, global average temperature is expected to warm at least twice as much as it has during the last 100 years.
- Ground-level air temperatures are expected to continue to warm more rapidly over land than oceans.
- Some parts of the world are projected to see larger temperature increases than the global average.

Future Precipitation and Storm Events

Patterns of precipitation and storm events, including both rain and snowfall are likely to change. However, some of these changes are less certain than the changes associated with temperature. Projections show that future precipitation and storm changes will vary by season and region. Some regions may have less precipitation, some may have more precipitation, and some may have little or no change. The amount of rain falling in heavy precipitation events is likely to increase in most regions, while storm tracks are projected to shift towards the poles. Climate models project the following precipitation and storm changes:

- Global average annual precipitation through the end of the century is expected to increase, although changes in the amount and intensity of precipitation will vary by region.
- The intensity of precipitation events will likely increase on average. This will be particularly pronounced in tropical and high-latitude regions, which are also expected to experience overall increases in precipitation.
- The strength of the winds associated with tropical storms is likely to increase. The amount of precipitation falling in tropical storms is also likely to increase.
- Annual average precipitation is projected to increase in some areas and decrease in others.

Future Ice, Snowpack, and Permafrost

Arctic sea ice is already declining drastically. The area of snow cover in the Northern Hemisphere has decreased since 1970. Permafrost temperature has increased over the last century, making it more susceptible to thawing. Over the next century, it is expected that sea ice will continue to decline, glaciers will continue to shrink, snow cover will continue to decrease, and permafrost will continue to thaw.

For every 2°F of warming, models project about a 15% decrease in the extent of annually averaged sea ice and a 25% decrease in September Arctic sea ice. The coastal sections of the Greenland and Antarctic ice sheets are expected to continue to melt or slide into the ocean. If the rate of this ice melting increases in the 21st century, the ice sheets could add significantly to global sea level rise. Glaciers are expected to continue to decrease in size. The rate of melting is expected to continue to increase, which will contribute to sea level rise.

Future Sea Level Change

Warming temperatures contribute to sea level rise by expanding ocean water, melting mountain glaciers and ice caps, and causing portions of the Greenland and Antarctic ice sheets to melt or flow into the ocean. Since 1870, global sea level has risen by about 8 inches. Estimates of future sea level rise vary for different regions, but global sea level for the next century is expected to rise at a greater rate than during the past 50 years. The contribution of thermal expansion, ice caps, and small glaciers to sea level rise is relatively well-studied, but the impacts of climate change on ice sheets are less understood and represent an active area of research. Thus, it is more difficult to predict how much changes in ice sheets will contribute to sea level rise. Greenland and Antarctic ice sheets could contribute an additional 1 foot of sea level rise, depending on how the ice sheets respond.

Regional and local factors will influence future relative sea level rise for specific coastlines around the world (figure $7.3.1.i$). For example, relative sea level rise depends on land elevation changes that occur as a result of subsidence (sinking) or uplift (rising), in addition to things such as local currents, winds, salinity, water temperatures, and proximity to thinning ice sheets. Assuming that these historical geological forces continue, a 2-foot rise in global sea level by 2100 would result in the following relative sea level rise:

- 2.3 feet at New York City
- 2.9 feet at Hampton Roads, Virginia
- 3.5 feet at Galveston, Texas
- 1 foot at Neah Bay in Washington state



Figure 7.3.1.*i*: The United States government paid for the residents of Isle De Jean Charles, an island south of Louisiana (that is also part of Louisiana), to relocate when it became inhabitable due to sea level rise. Image by Karen Apricot (CC-BY-SA).

Future Ocean Acidification

Ocean acidification is the process of ocean waters decreasing in pH. Oceans become more acidic as carbon dioxide (CO_2) emissions in the atmosphere dissolve in the ocean. This change is measured on the pH scale, with lower values being more acidic. The pH level of the oceans has decreased by approximately 0.1 pH units since pre-industrial times, which is equivalent to a 25% increase in acidity. The pH level of the oceans is projected to decrease even more by the end of the century as CO_2 concentrations are expected to increase for the foreseeable future. Ocean acidification adversely affects many marine species, including plankton, mollusks, shellfish, and corals. As ocean acidification increases, the availability of calcium carbonate will decline. Calcium carbonate is a key building block for the shells and skeletons of many marine organisms. If atmospheric CO_2 concentrations double, coral calcification rates are projected to decline by more than 30%. If CO_2 concentrations continue to rise at their current rate, corals could become rare on tropical and subtropical reefs by 2050.

Mismatched Interactions

Climate change also affects phenology, the study of the effects of climatic conditions on the timing of periodic lifecycle events, such as flowering in plants or migration in birds. Researchers have shown that 385 plant species in Great Britain are flowering 4.5 days sooner than was recorded earlier during the previous 40 years. In addition, insect-pollinated species were more likely to flower earlier than wind-pollinated species. The impact of changes in flowering date would be mitigated if the insect pollinators emerged earlier. This mismatched timing of plants and pollinators could result in injurious ecosystem effects because, for continued survival, insect-pollinated plants must flower when their pollinators are present.

Likewise, migratory birds rely on daylength cues, which are not influenced by climate change. Their insect food sources, however, emerge earlier in the year in response to warmer temperatures. As a result, climate change decreases food availability for migratory bird species.

Spread of Disease

This rise in global temperatures will increase the range of disease-carrying insects and the viruses and pathogenic parasites they harbor. Thus, diseases will spread to new regions of the globe. This spread has already been documented with dengue fever, a disease the affects hundreds of millions per year, according to the World Health Organization. Colder temperatures typically limit the distribution of certain species, such as the mosquitoes that transmit malaria, because freezing temperatures destroy their eggs.

Not only will the range of some disease-causing insects expand, the increasing temperatures will also accelerate their lifecycles, which allows them to breed and multiply quicker, and perhaps evolve pesticide resistance faster. In addition to dengue fever, other diseases are expected to spread to new portions of the world as the global climate warms. These include malaria, yellow fever, West Nile virus, zika virus, and chikungunya.

Climate change does not only increase the spread of diseases in humans. Rising temperatures are associated with greater amphibian mortality due to chytridiomycosis (see Invasive Species). Similarly, warmer temperatures have exacerbated bark beetle infestations of coniferous trees, such as pine an spruce.

Climate Change Affects Everyone

Our lives are connected to the climate**.** Human societies have adapted to the relatively stable climate we have enjoyed since the last ice age which ended several thousand years ago. A warming climate will bring changes that can affect our water supplies, agriculture, power and transportation systems, the natural environment, and even our own health and safety.

Carbon dioxide can stay in the atmosphere for nearly a century, on average, so Earth will continue to warm in the coming decades. The warmer it gets, the greater the risk for more severe changes to the climate and Earth's system. Although it's difficult to predict the exact impacts of climate change, what's clear is that the climate we are accustomed to is no longer a reliable guide for what to expect in the future.

We can reduce the risks we will face from climate change**.** By making choices that reduce greenhouse gas pollution, and preparing for the changes that are already underway, we can reduce risks from climate change. Our decisions today will shape the world our children and grandchildren will live in.

You can take steps at home, on the road, and in your office to reduce greenhouse gas emissions and the risks associated with climate change. Many of these steps can save you money. Some, such as walking or biking to work, can even improve your health! You can also get involved on a local or state level to support energy efficiency, clean energy programs, or other climate programs.

Suggested Supplementary Reading

Intergovernmental Panel on Climate Change. 2013. 5th Assessment: Summary for Policymakers.

NASA. 2018. Global Climate Change: Vital Signs of the Planet. This website by NASA provides a multi-media smorgasbord of engaging content. Learn about climate change using data collected by NASA satellites and more.

Attributions

Modified by Melissa Ha from the following sources:

- Climate and the Effects of Global Climate Change from *General Biology* by OpenStax (licensed under CC-BY)
- Climate Change from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)
- Carbon Cycle from *Biology* by John W. Kimball (licensed under CC-BY)

This page titled 7.3.1: The Greenhouse Effect and Climate Change is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **21.1: The Greenhouse Effect and Climate Change** by Melissa Ha and Rachel Schleiger is licensed CC BY-SA 4.0.

7.3.2: Data Dive- Flooding in Venice Italy

Overview

Historically, Venice Italy has a history of flooding that is caused from a multitude of both natural and human caused phenomena. Currently, these issues are being exacerbated by sea level changes as a result of climate change. As such, the City of Venice has been carefully tracking frequency of flooding through the years. The graph below represents the frequency of water level was above 110cm per 10-year period:

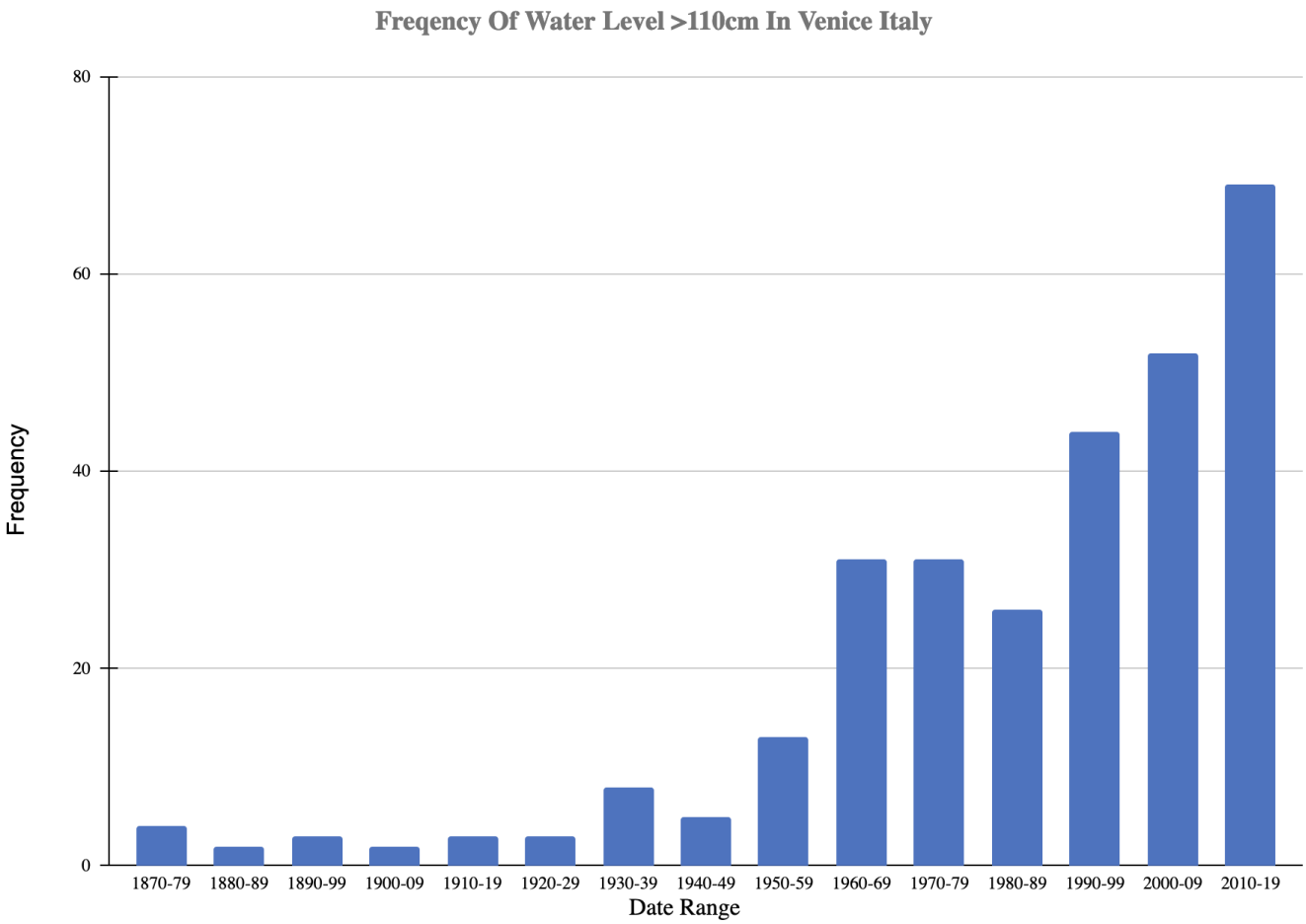


Figure 7.3.2.*a*: Frequency of flooding exceeding 110cm in Venice Italy. Graph by Rachel Schleiger (CC-BY-NC) modified from data published by the City of Venice.

Questions

1. What is the independent (explanatory) variable and the dependent (response) variable?
2. What question(s) are the authors trying to answer with this graph?
3. What are the results of this graph illustrating about flooding in Venice?
4. How can other coastal cities around the world learn from what patterns are emerging in this graph?
5. Provide a brief reflection for how these results make you feel and what they make you think about.

Raw Data For Above Graph(s)

Table 7.3.2.*a***:** Raw data for the frequency of flooding exceeding 110cm in Venice Italy. Graph by Rachel Schleiger (CC-BY-NC) modified from data published by the City of Venice.

Ten Year Interval	Frequency
1870-79	4
1880-89	2
1890-99	3
1900-09	2
1910-19	3
1920-29	3
1930-39	8
1940-49	5
1950-59	13
1960-69	31
1970-79	31
1980-89	26
1990-99	44
2000-09	52
2010-19	69

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 7.3.2: Data Dive- Flooding in Venice Italy is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **21.2: Data Dive- Flooding in Venice Italy** by Melissa Ha and Rachel Schleiger has no license indicated.

7.3.3: Review

Summary

After completing this chapter you should be able to...

- Define global climate change.
- Summarize the effects of the Industrial Revolution on global atmospheric carbon dioxide concentration.
- Describe three natural factors affecting long-term global climate.
- List two or more greenhouse gases and describe their role in the greenhouse effect.

The historical record shows that the climate system varies naturally over a wide range of time scales. In general, climate changes prior to the **Industrial Revolution** in the 1700s can be explained by natural causes, such as changes in solar energy, volcanic eruptions, and natural changes in greenhouse gas concentrations. Recent climate changes, however, cannot be explained by natural causes alone. Natural causes are very unlikely to explain most observed warming, especially warming since the mid-20th century. Rather, human activities can explain most of that warming.

The primary human activity affecting the amount and rate of climate change is **greenhouse gas** emissions from the burning of fossil fuels. Greenhouse gas concentrations in the atmosphere will continue to increase unless the billions of tons of our annual emissions decrease substantially. Increased concentrations are expected to increase Earth's average temperature, influence the patterns and amounts of precipitation, reduce ice and snow cover, as well as **permafrost**, raise sea level and increase the acidity of the oceans. These changes will impact our food supply, water resources, infrastructure, ecosystems, and even our own health.

Attribution

Modified by Melissa Ha from Air Pollution, Climate Change, & Ozone Depletion from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 7.3.3: Review is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **21.3: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY 4.0.

CHAPTER OVERVIEW

8: Sustainability

Connecting Environmental Science to Sustainability

Both environmental science and sustainability originated out of necessity due to past and current issues to find solutions for the now and the future. As such, there are many parallels to these subjects basic objectives. However, environmental science focuses on environmental/ecological solutions (as outlined in this book), whereas sustainability focuses on environmental/ecological, social, and economic goals. That's even more interdisciplinary, and thus complex, than environmental science! As such, this chapter will solely introduce essential concepts of sustainability and will not be entirely comprehensive.



Figure 8.*a*: Example goals of sustainability. Image by Wikimedia Commons (CC-BY-SA4.0)

8.1: Introduction to Sustainability
8.2: History of Sustainability
8.3: Environmental Ethics
8.3.1: Environmental Justice and Indigenous Struggles
8.3.2: Perspectives of the Environment
8.4: Environmental Policy
8.5: Sustainable Cities
8.6: Sustainability and the Future
8.7: Data Dive- Sustainable Investments
8.8: Review

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 8: Sustainability is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

8.1: Introduction to Sustainability

Chapter Hook

Cities in today's modern world are struggling to balance a sustainable society, economy, and environment. According to a 2016 report that rated sustainability across 32 indicators, Zurich, Switzerland, was ranked number one. Zurich invests in renewable energies, sustainable public transport, public green space, and public education. One of the most notable ways Zurich is leading global sustainability efforts is their dedication to keeping their carbon dioxide emissions low. Figure 8.1.*a* below shows the difference in emissions for the United States and Switzerland:

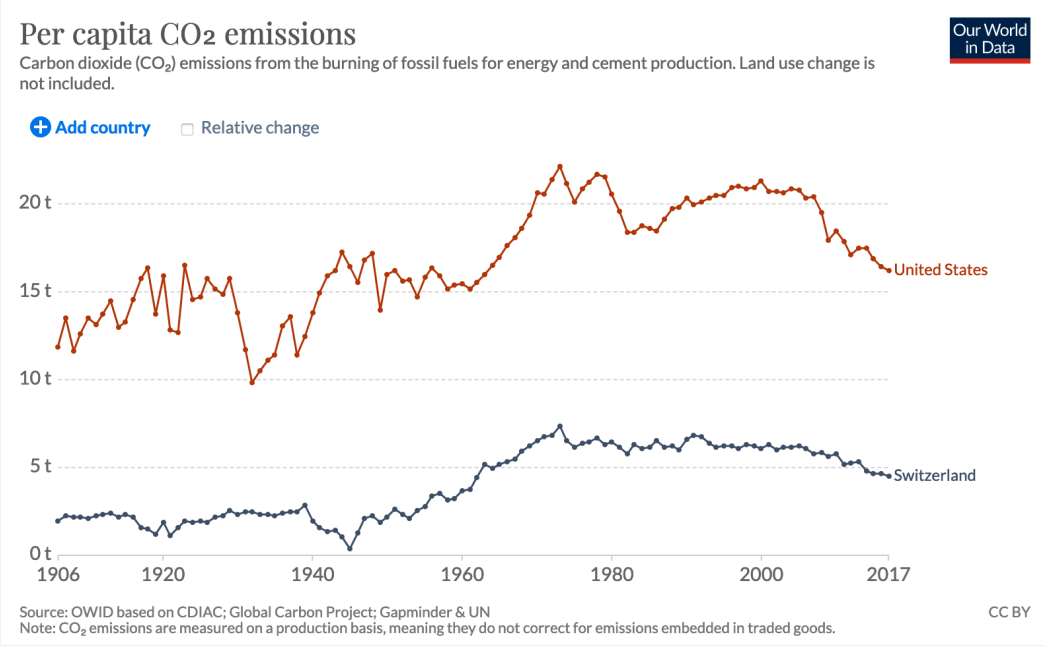


Figure 8.1.*a*: Per capita emissions of carbon dioxide between the United States and Switzerland. Graph by Our World in Data (CC-BY4.0)

Introduction

Sustainability is derived from two Latin words: *sus* which means up, and *tenere* which means to hold. Thus, sustainability is essentially about holding up human existence by meeting the needs of the present without compromising the ability of future generations to meet their needs.

There are three dimensions that sustainability seeks to integrate: economic, environmental, and social (including sociopolitical).

- Economic interests define the framework for making decisions, the flow of financial capital, and the facilitation of commerce, including the knowledge, skills, competences and other attributes embodied in individuals that are relevant to economic activity.
- Environmental aspects recognize the diversity and interdependence within living systems, the goods and services produced by the world's ecosystems, and the impacts of human wastes.
- Social/Socio-political refers to interactions between institutions/firms and people, functions expressive of human values, aspirations and well-being, ethical issues, and decision-making that depends upon collective action.

The intersection of social and economic elements can form the basis of social "equitability". In the sense of enlightened management, "viability" is formed through consideration of economic and environmental interests. Between environment and social elements lies "bearability," the recognition that the functioning of societies is dependent on environmental resources and services. At the intersection of all three of these lies sustainability (figure 8.1.*b*).

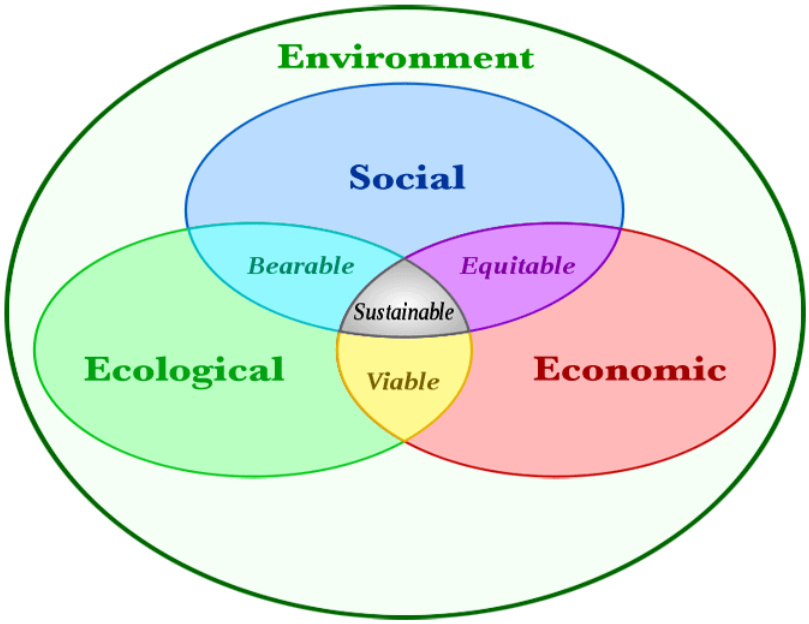


Figure 8.1.*b*: Venn diagram of defining the components of sustainability. Image by Andrew, Sunray, based on "File:Sustainable development.svg" by Johann Dréo in Wikimedia Commons (CC-BY-SA4.0)

The three main elements of the sustainability paradigm are thought of as equally important, however, tradeoffs occur depending on the local/global objective. For example, in some instances it may be deemed necessary to degrade a particular ecosystem in order to facilitate commerce, or food production, or housing. In reality, the extent to which tradeoffs can be made before irreversible damage results is not always known, and in any case there are definite limits on how much substitution among the three elements is wise (to date, humans have treated economic development as the dominant one of the three). This has led to the notion of strong sustainability, where tradeoffs among natural, human, and social capital are not allowed or are very restricted, and weak sustainability, where tradeoffs are unrestricted or have few limits. Whether or not one follows the strong or weak form of sustainability, it is important to understand that while economic and social systems are human creations, the environment is not. Rather, a functioning environment underpins both society and the economy.

Attribution

Modified by Rachel Schleiger from Sustainability: A Comprehensive Foundation by Openstax (licensed under CC-BY)

This page titled 8.1: Introduction to Sustainability is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **22.1: Introduction to Sustainability** by Melissa Ha and Rachel Schleiger has no license indicated.

8.2: History of Sustainability

Taking The Long View: Sustainability an Evolutionary and Ecological Perspective

Of the different forms of life that have inhabited the Earth in its three to four billion year history, 99.9% are now extinct. Against this backdrop, the human enterprise with its roughly 200,000-year history barely merits attention. As the American novelist Mark Twain once remarked, if our planet's history were to be compared to the Eiffel Tower, human history would be a mere smear on the very tip of the tower. But while modern humans (*Homo sapiens*) might be insignificant in geologic time, we are by no means insignificant in terms of our recent planetary impact. A 1986 study estimated that 40% of the product of terrestrial plant photosynthesis — the basis of the food chain for most animal and bird life — was being appropriated by humans for their use. More recent studies estimate that 25% of photosynthesis on continental shelves (coastal areas) is ultimately being used to satisfy human demand. Human appropriation of such natural resources is having a profound impact upon the wide diversity of other species that also depend on them.

Evolution normally results in the generation of new lifeforms at a rate that outstrips the extinction of other species; this results in strong biological diversity. However, scientists have evidence that, for the first observable time in evolutionary history, another species — *Homo sapiens* — has upset this balance to the degree that the rate of species extinction is now estimated at 10,000 times the rate of species renewal. Human beings, just one species among millions, are crowding out the other species we share the planet with. Evidence of human interference with the natural world is visible in practically every ecosystem from the presence of pollutants in the stratosphere to the artificially changed courses of the majority of river systems on the planet. It is argued that ever since we abandoned nomadic, gatherer-hunter ways of life for settled societies some 12,000 years ago, humans have continually manipulated their natural world to meet their needs. While this observation is a correct one, the rate, scale, and the nature of human-induced global change — particularly in the post-industrial period — is unprecedented in the history of life on Earth.

There are three primary reasons for this:

1. Mechanization of both industry and agriculture in the last century resulted in vastly improved labor productivity which enabled the creation of goods and services. Since then, scientific advance and technological innovation — powered by ever-increasing inputs of fossil fuels and their derivatives — have revolutionized every industry and created many new ones. The subsequent development of western consumer culture, and the satisfaction of the accompanying disposable mentality, has generated material flows of an unprecedented scale. The Wuppertal Institute estimates that humans are now responsible for moving greater amounts of matter across the planet than all natural occurrences (earthquakes, storms, etc.) put together.
2. The sheer size of the human population is unprecedented. Every passing year adds another 90 million people to the planet. Even though the environmental impact varies significantly between countries (and within them), the exponential growth in human numbers, coupled with rising material expectations in a world of limited resources, has catapulted the issue of distribution to prominence. Global inequalities in resource consumption and purchasing power mark the clearest dividing line between the haves and the have-nots. It has become apparent that present patterns of production and consumption are unsustainable for a global population that is projected to reach between 12 billion by the year 2050. If ecological crises and rising social conflict are to be countered, the present rates of over-consumption by a rich minority, and under-consumption by a large majority, will have to be brought into balance.
3. It is not only the rate and the scale of change but the nature of that change that is unprecedented. Human inventiveness has introduced chemicals and materials into the environment which either do not occur naturally at all, or do not occur in the ratios in which we have introduced them. These persistent chemical pollutants are believed to be causing alterations in the environment, the effects of which are only slowly manifesting themselves, and the full scale of which is beyond calculation. CFCs and PCBs are but two examples of the approximately 100,000 chemicals currently in global circulation. (Between 500 and 1,000 new chemicals are being added to this list annually.) The majority of these chemicals have not been tested for their toxicity on humans and other life forms, let alone tested for their effects in combination with other chemicals. These issues are now the subject of special UN and other intergovernmental working groups.

The Evolution of Sustainability Itself

Our Common Future (1987), the report of the World Commission on Environment and Development, is widely credited with having popularized the concept of **sustainable development.** It defines sustainable development in the following ways…

- …development that meets the needs of the present without compromising the ability of future generations to meet their own needs.
- … sustainable development is not a fixed state of harmony, but rather a process of change in which the exploitation of resources, the orientation of the technological development, and institutional change are made consistent with future as well as present needs.



Figure 8.2.*a***:** Cover of Our Common Future. Image by Sigurður Kaiser in Wikimedia Commons (CC-BY-SA4.0)

The concept of sustainability, however, can be traced back much farther to the oral histories of indigenous cultures. For example, the principle of inter-generational equity is captured in the Inuit saying, 'we do not inherit the Earth from our parents, we borrow it from our children'. The Native American 'Law of the Seventh Generation' is another illustration. According to this, before any major action was to be undertaken its potential consequences on the seventh generation had to be considered. For a species that at present is only 6,000 generations old and whose current political decision-makers operate on time scales of months, or few years at most, the thought that other human cultures have based their decision-making systems on time scales of many decades seems wise but unfortunately inconceivable in the current political climate.

Precautionary Principle

The **precautionary principle** is an important concept in environmental sustainability. A 1998 consensus statement characterized the precautionary principle this way: "when an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically". For example, if a new pesticide chemical is created, the precautionary principle would dictate that we presume, for the sake of safety, that the chemical may have potential negative consequences for the environment and/or human health, even if such consequences have not been proven yet. In other words, it is best to proceed cautiously in the face of incomplete knowledge about something's potential harm.

Attribution

Modified by Melissa Ha and Rachel Schleiger from Environment and Sustainability from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 8.2: History of Sustainability is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **22.2: History of Sustainability** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

SECTION OVERVIEW

8.3: Environmental Ethics

Introduction

The concept of **ethics** involves standards of conduct. These standards help to distinguish between behavior that is considered right and that which is considered wrong. As we all know, it is not always easy to distinguish between right and wrong, as there is no universal code of ethics. For example, a poor farmer clears an area of rainforest in order to grow crops. Some would not oppose this action, because the act allows the farmer to provide a livelihood for his family. Others would oppose the action, claiming that the deforestation will contribute to soil erosion and global warming. Right and wrong are usually determined by an individual's morals, and to change the ethics of an entire society, it is necessary to change the individual ethics of a majority of the people in that society.

Moral extensionism defines how far a persons values extend outside of themselves (Figure 7.2a). There are many variables that can influence where the limit lies for each individual person. For example: religion, culture, education, and personal interests (just to name a few).



Figure 8.3.*a***:** Scope of consideration for moral extensionism. Image by Rachel Schleiger (CC-BY-NC).

Attribution

Modified by Rachel Schleiger from Sustainability: A Comprehensive Foundation by Openstax (licensed under CC-BY)

Topic hierarchy

8.3.1: Environmental Justice and Indigenous Struggles

8.3.2: Perspectives of the Environment

This page titled 8.3: Environmental Ethics is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

8.3.1: Environmental Justice and Indigenous Struggles

Environmental Equity

Environmental equity describes a country, or world, in which no single group or community faces disadvantages in dealing with environmental hazards, disasters, or pollution. While much progress is being made to improve resource efficiency, far less progress has been made to improve resource distribution. Currently, just one-fifth of the global population is consuming three-quarters of the earth's resources.

Global Consumption Inequality - 24% of the global population (mostly in the high income countries) accounts for...

- 92% Cars
- 70% Carbon dioxide emissions
- 86% Copper and aluminum
- 81% Paper
- 80% Iron and steel
- 48% Cereal crops
- 60% Artificial fertilizer

If the remaining three-quarters were to exercise their right to grow to the level of the rich minority it would result in ecological devastation. So far, global income inequalities and lack of purchasing power have prevented poorer countries from reaching the standard of living (and also resource consumption/waste emission) of the industrialized countries. Countries such as China, Brazil, India, and Malaysia are, however, catching up fast. In such a situation, global consumption of resources and energy needs to be drastically reduced to a point where it can be repeated by future generations. But who will do the reducing? Poorer nations want to produce and consume more. Yet so do richer countries: their economies demand ever greater consumption-based expansion. Such stalemates have prevented any meaningful progress towards equitable and sustainable resource distribution at the international level. These issue of fairness and distributional justice remain unresolved.

Environmental Justice

Environmental justice is defined as the fair treatment and meaningful involvement of all people regardless of race, color, national origin, or income with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies. It will be achieved when everyone enjoys the same degree of protection from environmental and health hazards and equal access to the decision-making process to have a healthy environment in which to live, learn, and work.

In Flint Michigan, the city decided to save money by drawing water for residents from the Flint River in 2014. Residents complained about the taste, smell, and color of the water. A couple of scientific studies determined that there was a failure to apply corrosion inhibitors to the water, resulting in lead from aging pipes to leach into the water supply, exposing around 100,000 residents to elevated lead levels. Although the city switched back to it's original water source, the damage was already done. Not only did this crisis expose so many to lead, especially concerning for children, but also most pipes needed to be replaced as they were no longer safe to use (Figure a). In 2017 a settlement was reached to replace all the pipes, a $87 million project. However, as of Jan 2021 this project is only just now approaching completion. Citywide lead levels have tested within the safe range for consumption, though houses considered high risk do continue to show elevated levels.



Figure 8.3.1.*a***:** Flint residents trying to bring awareness of Flint Water Crisis through peaceful demonstration. Image by Shannon Nobles in Wikimedia Commons (CC-BY-SA-4.0).

During the 1980's marginalized groups protested that hazardous waste sites were preferentially sited in their neighborhoods. In 1987, Benjamin Chavis of the United Church of Christ Commission for Racism and Justice coined the term environmental racism to describe such a practice. **Environmental racism** occurs when marginalized groups and people of color are disproportionately exposed to environmental hazards.. Confounding this issue, most hazardous waste sites are located on property that was used as disposal sites long before modern facilities and disposal methods were available. Areas around such sites are typically depressed economically, often as a result of past disposal activities. Persons with low incomes are often financially constrained to live in these areas.

Decisions in citing hazardous waste facilities are generally made on the basis of economics, geological suitability and the political climate. For example, a site must have a soil type and geological profile that prevents hazardous materials from moving into local aquifers. The cost of land is also an important consideration. The high cost of buying land would make it economically unfeasible to build a hazardous waste site in Beverly Hills. Some communities have seen a hazardous waste facility as a way of improving their local economy and quality of life. Emelle County, Alabama had illiteracy and infant mortality rates that were among the highest in the nation. A landfill constructed there provided jobs and revenue that ultimately helped to reduce both figures.

In an ideal world, there would be no hazardous wastes to plague this planet. Unfortunately, we live in a world with rampant pollution, dumping of hazardous wastes, and people with "living for the now" attitudes. Our industrialized society has necessarily produced wastes during the manufacture of products for our basic needs. Until technology can find a way to manage (or eliminate) hazardous waste, disposal facilities will be necessary to protect both humans and the environment. By the same token, this problem must be addressed. Industry and society must become more socially sensitive in the selection of future hazardous waste sites. All humans who help produce hazardous wastes must share the burden of dealing with those wastes, not just the marginalized and low income groups.

Indigenous People

Since the end of the 15th century, much of the world has been claimed and colonized by established nations. Invariably, these conquered frontiers were home to people indigenous to those regions. Some were wiped out or assimilated by the invaders, while others survived while trying to maintain their unique cultures and way of life. The United Nations officially classifies **indigenous**

people as those “having an historical continuity with pre-invasion and pre-colonial societies,” and “consider themselves distinct from other sectors of the societies now prevailing in those territories or parts of them.” Furthermore, indigenous people are “determined to preserve, develop and transmit to future generations, their ancestral territories, and their ethnic identity, as the basis of their continued existence as peoples in accordance with their own cultural patterns, social institutions and legal systems.” A few of the many groups of indigenous people around the world are: the many tribes of Native Americans (i.e., Navajo, Sioux) in the contiguous 48 states, the Inuit of the arctic region from Siberia to Canada, the rainforest tribes in Brazil, and the Ainu of northern Japan.

Many problems face indigenous people including the lack of human rights, exploitation of their traditional lands and themselves, and degradation of their culture. In response to the problems faced by these people, the United Nations proclaimed an “International Decade of the World’s Indigenous People” beginning in 1994. The main objective of this proclamation, according to the United Nations, is “the strengthening of international cooperation for the solution of problems faced by indigenous people in such areas as human rights, the environment, development, health, culture and education.” Its major goal is to protect the rights of indigenous people. Such protection would enable them to retain their cultural identity, such as their language and social customs, while participating in the political, economic and social activities of the region in which they reside.

Despite the lofty U.N. goals, the rights and feelings of indigenous people are often ignored or minimized, even by supposedly culturally sensitive developed countries. In the United States many of those in the federal government are pushing to exploit oil resources in the Arctic National Wildlife Refuge on the northern coast of Alaska. The “Gwich’in,” an indigenous people who rely culturally and spiritually on the herds of caribou that live in the region, claim that drilling in the region would devastate their way of life (Figure b). Thousands of years of culture would be destroyed for a few months’ supply of oil. Drilling efforts have been stymied in the past, but mostly out of concern for environmental factors and not necessarily the needs of the indigenous people. Curiously, another group of indigenous people, the “Inupiat Eskimo,” favor oil drilling in the Arctic National Wildlife Refuge. Because they own considerable amounts of land adjacent to the refuge, they would potentially reap economic benefits from the development of the region.



Figure 8.3.1.*b*: Bernadette Demietieff speaking in support of protect the Arctic to protect Gwich'in land from oil drilling, Alaska, c. 2019. Image by: Deb Haaland in Wikimedia Commons (CC-BY-SA-4.0).

The heart of most environmental conflicts faced by governments usually involves what constitutes proper and sustainable levels of development. For many indigenous peoples, sustainable development constitutes an integrated wholeness, where no single action is separate from others. They believe that sustainable development requires the maintenance and continuity of life, from generation to generation and that humans are not isolated entities, but are part of larger communities, which include the seas, rivers, mountains, trees, fish, animals and ancestral spirits. These, along with the sun, moon and cosmos, constitute a whole. From the point of view of

indigenous people, sustainable development is a process that must integrate spiritual, cultural, economic, social, political, territorial and philosophical ideals.

Attribution

Modified by Melissa Ha and Rachel Schleiger from Environmental Justice and Indigenous Struggles and Environment and Sustainability from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 8.3.1: Environmental Justice and Indigenous Struggles is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **22.3.1: Environmental Justice and Indigenous Struggles** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

8.3.2: Perspectives of the Environment

Frontier Ethic

The ways in which humans interact with the land and its natural resources are determined by ethical attitudes and behaviors. Early European settlers in North America rapidly consumed the natural resources of the land. After they depleted one area, they moved westward to new frontiers. Their attitude towards the land was that of a **frontier ethic**. A frontier ethic assumes that the earth has an unlimited supply of resources. If resources run out in one area, more can be found elsewhere or alternatively human ingenuity will find substitutes. This attitude sees humans as masters who manage the planet. The frontier ethic is completely anthropocentric (human-centered), for only the needs of humans are considered.

Most industrialized societies experience population and economic growth that are based upon this frontier ethic, assuming that infinite resources exist to support continued growth indefinitely. In fact, economic growth is considered a measure of how well a society is doing. The late economist Julian Simon pointed out that life on Earth has never been better, and that population growth means more creative minds to solve future problems and give us an even better standard of living. However, now that the human population has passed seven billion and few frontiers are left, many are beginning to question the frontier ethic. Such people are moving toward an environmental ethic, which includes humans as part of the natural community rather than managers of it. Such an ethic places limits on human activities (e.g., uncontrolled resource use), that may adversely affect the natural community.

Some of those still subscribing to the frontier ethic suggest that outer space may be the new frontier. If we run out of resources (or space) on earth, they argue, we can simply populate other planets. This seems an unlikely solution, as even the most aggressive colonization plan would be incapable of transferring people to extraterrestrial colonies at a significant rate. Natural population growth on earth would outpace the colonization effort. A more likely scenario would be that space could provide the resources (e.g. from asteroid mining) that might help to sustain human existence on earth.

Sustainable Ethic

A **sustainable ethic** is an environmental ethic by which people treat the earth as if its resources are limited. This ethic assumes that the earth's resources are not unlimited and that humans must use and conserve resources in a manner that allows their continued use in the future. A sustainable ethic also assumes that humans are a part of the natural environment and that we suffer when the health of a natural ecosystem is impaired. A sustainable ethic includes the following tenets:

- The earth has a limited supply of resources.
- Humans must conserve resources.
- Humans share the earth's resources with other living things.
- Growth is not sustainable.
- Humans are a part of nature.
- Humans are affected by natural laws.
- Humans succeed best when they maintain the integrity of natural processes sand cooperate with nature.

For example, if a fuel shortage occurs, how can the problem be solved in a way that is consistent with a sustainable ethic? The solutions might include finding new ways to conserve oil or developing renewable energy alternatives. A sustainable ethic attitude in the face of such a problem would be that if drilling for oil damages the ecosystem, then that damage will affect the human population as well. A sustainable ethic can be either anthropocentric or biocentric (life-centered). An advocate for conserving oil resources may consider all oil resources as the property of humans. Using oil resources wisely so that future generations have access to them is an attitude consistent with an anthropocentric ethic. Using resources wisely to prevent ecological damage is in accord with a biocentric ethic.

Land Ethic

Aldo Leopold, an American wildlife natural historian and philosopher, advocated a **biocentric ethic** in his book, *A Sand County Almanac*. He suggested that humans had always considered land as property, just as ancient Greeks considered slaves as property. He believed that mistreatment of land (or of slaves) makes little economic or moral sense, much as today the concept of slavery is

considered immoral. All humans are merely one component of an ethical framework. Leopold suggested that land be included in an ethical framework, calling this the land ethic.

“The **land ethic** simply enlarges the boundary of the community to include soils, waters, plants and animals; or collectively, the land. In short, a land ethic changes the role of *Homo sapiens* from conqueror of the land-community to plain member and citizen of it. It implies respect for his fellow members, and also respect for the community as such.” (Aldo Leopold, 1949)

Leopold divided conservationists into two groups: one group that regards the soil as a commodity and the other that regards the land as biota, with a broad interpretation of its function. If we apply this idea to the field of forestry, the first group of conservationists would grow trees like cabbages, while the second group would strive to maintain a natural ecosystem. Leopold maintained that the conservation movement must be based upon more than just economic necessity. Species with no discernible economic value to humans may be an integral part of a functioning ecosystem. The land ethic respects all parts of the natural world regardless of their utility, and decisions based upon that ethic result in more stable biological communities.

“Anything is right when it tends to preserve the integrity, stability and beauty of the biotic community. It is wrong when it tends to do otherwise.” (Aldo Leopold, 1949)

Case Study: Hetch Hetchy

In 1913, the Hetch Hetchy Valley – located in Yosemite National Park in California – was the site of a conflict between two factions, one with an anthropocentric ethic and the other, a biocentric ethic. As the last American frontiers were settled, the rate of forest destruction started to concern the public.



Figure 8.3.2.*a***:** Aerial view of Hetch Hetchy Reservoir, California, USA. Image by Ken Lund in Wikimedia Commons (CC-BY-SA)

The conservation movement gained momentum, but quickly broke into two factions. One faction, led by Gifford Pinchot, Chief Forester under Teddy Roosevelt, advocated utilitarian conservation (i.e., conservation of resources for the good of the public). The other faction, led by John Muir, advocated preservation of forests and other wilderness for their inherent value. Both groups rejected the first tenet of frontier ethics, the assumption that resources are limitless. However, the conservationists agreed with the rest of the tenets of frontier ethics, while the preservationists agreed with the tenets of the sustainable ethic.

The Hetch Hetchy Valley was part of a protected National Park, but after the devastating fires of the 1906 San Francisco earthquake, residents of San Francisco wanted to dam the valley to provide their city with a stable supply of water. Gifford Pinchot favored the dam.

> *"As to my attitude regarding the proposed use of Hetch Hetchy by the city of San Francisco…I am fully persuaded that… the injury…by substituting a lake for the present swampy floor of the valley…is altogether unimportant compared with the benefits to be derived from it's use as a reservoir.*
>
> *"The fundamental principle of the whole conservation policy is that of use, to take every part of the land and its resources and put it to that use in which it will serve the most people." (Gifford Pinchot, 1913)*

John Muir, the founder of the Sierra Club and a great lover of wilderness, led the fight against the dam. He saw wilderness as having an intrinsic value, separate from its utilitarian value to people. He advocated preservation of wild places for their inherent beauty and for the sake of the creatures that live there. The issue aroused the American public, who were becoming increasingly alarmed at the growth of cities and the destruction of the landscape for the sake of commercial enterprises. Key senators received thousands of letters of protest.

> *"These temple destroyers, devotees of ravaging commercialism, seem to have a perfect contempt for Nature, and instead of lifting their eyes to the God of the Mountains, lift them to the Almighty Dollar." (John Muir, 1912)*

Despite public protest, Congress voted to dam the valley. The preservationists lost the fight for the Hetch Hetchy Valley, but their questioning of traditional American values had some lasting effects. In 1916, Congress passed the "National Park System Organic Act," which declared that parks were to be maintained in a manner that left them unimpaired for future generations. As we use our public lands, we continue to debate whether we should be guided by preservationism or conservationism.

Attribution

Modified by Melissa Ha and Rachel Schleiger from Environmental Ethics from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY)

This page titled 8.3.2: Perspectives of the Environment is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **22.3.2: Perspectives of the Environment** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

8.4: Environmental Policy

Introduction

It is not uncommon to think of the sustainability paradigm as being a recent interpretation of environmental policy, one that was given credence by the United Nations report "Our Common Future" (the Brundtland Report) when it was first presented in 1987. However, it would be a mistake to conclude that sustainability as a mental construct and policy framework for envisioning the relationship of humans and nature came into being suddenly and at a single moment in time. Most environmental historians who have studied U.S. policy have discerned at least three distinct periods during which new concepts and ideas, scientific understandings, technological advances, political institutions, and laws and regulations came or were brought into being in order to understand and manage human impacts on the environment.

1. **The American conservation movement:** This 19th century movement had the main objective for conservation of American land and resources, it grew from opposite philosophies. On one side you had people like Gifford Pinchot (first head of the US National Forest Service) that were very pragmatic about their desire for conservation. For example, these people observed that policies for harvesting some resource (like lumber) were not sustainable, so they began conservation/sustainability research and new policies so the resource would not run out in the foreseeable future. On the other side were people like John Muir and Theodore Roosevelt that saw land/habitats/wildlife as intrinsic entities, meaning that had value in their own right (meaning nothing to do with money).
2. **The rise of environmental risk management as a basis for policy:** The beginnings of environmental risk management can be traced to the fields of public health, industrial hygiene, and sanitary engineering, which came into prominence in the latter decades of the 19th century and beginning of the 20th. The spread of disease was a particularly troublesome problem as the country continued to urbanize. In addition, environmental scientists of the day were alarmed by the extent and degree of damage that they were documenting. The publication of *Silent Spring* in 1962 by Rachel Carson (1907-1964), about the impact of the widespread and indiscriminate use of pesticides, was a watershed moment, bringing environmental concerns before a large portion of the American, and global, public. Carson, collected scientifically documented evidence on the effects of pesticides, particularly DDT, heptachlor, and dieldrin, on humans and mammals, and the systemic disruption they caused to ecosystems. *Silent Spring* is credited with bringing about a ban on the use of DDT in the United States, and setting in motion a chain of events that would ultimately result in the transformation of environmental public policy from one based on the problems and attitudes that brought about nineteenth century conservation, to one based on the management of risks from chemical toxins. The U.S. Environmental Protection Agency was established in 1970, just eight years after the publication of *Silent Spring*. The same year Earth Day was created.
3. **The integration of social and economic factors to create what we now refer to as the sustainability paradigm:** Eventually, we (humanity) realized that in order to solve environmental issues (like ones identified by Rachel Carson), that the "human factors" (both social and economical) had to be accounted for. This is what defines the sustainability paradigm: environmental, social, and economic disciplines. Only Band-Aid solutions can be applied if we only take a one-minded approach.

History of Environmental Policy

Environmental policies are typically driven by problems of the day, real and perceived, that require systemic solutions. For example:

- Early conservationists were alarmed at the inefficiencies of human resource management and the encroachment of humans on unspoiled lands... Leading to conservation legislation.
- During the 20th century many groups (scientists, economists, politicians, and ordinary citizens) became alarmed and fearful of the consequences of toxic pollutant loads to the environment that included localized effects on human health and well-being... Leading to clean air and water legislation.
- As we proceed into the 21st century, an array of complex problems that have the potential to alter substantially the structure and well-being of large segments of human societies, calls for a renewal and reassessment of our approach to environmental policy.

This has, thus far, proven to be a difficult transition. Many of these complex problems have multiple causes and impacts, affect some groups of people more than others, are economically demanding, and are often not as visibly apparent to casual observers as previous impacts, nor are the benefits perceived to be commensurate with costs. Devising a regulatory strategy for such problems requires an adaptive and flexible approach that current laws do not foster.

Table 8.4.*a***:** Table illustrating evolution of environmental policy. Table by Sustainability: A Comprehensive Foundation in Openstax (CC-BY).

	1850-1920	1960-1990	1990-Present
Focus	Conservation/sanitation	Media/site/problem specific	Complex regional/ global problems
Outcome	Land preservation/efficiency/control of disease	Manage anthropocentric and ecological risk	Global sustainable development
Principal Activity	Resource management reform/simple contaminant controls	Compliance/remediation/technological emphasis on problem solving	Integration of social, economic, and technological information for holistic problem solving
Economic Focus	Profit maximization/public health	Cost minimization	Strategic investments/long-term societal well-being
Regulatory Activity	Low	Heavy	Adaptive and flexible
Conceptual Model	Expansion vs preservation	Command and control	Systems/life cycle approach
Disciplinary Approach	Disciplinary and insular	Multidisciplinary	Interdisciplinary/integrative

Table 8.4.*b***:** Table illustrating major environmental legislation, society, and organization conceptions. Table by Rachel Schleiger (CC-BY-NC).

Year	Policy/Society/Organization Conception
1899	Refuse Act
1905	National Forest Service
1916	National Park Service
1918	Migratory Bird Treaty Act
1930's	State Parks
1948	Federal Water Pollution Control Act
1950	The Nature Conservancy
1955	Air Pollution Control Act
1961	World Wildlife Federation
1963	Clean Air Act (1970, 1977, 1990 updates)
1965	Solid Waste Disposal Act (1976 update)
1965	Water Quality Act (1987 update)
1967	Air Quality Act
1969	National Environmental Policy Act
1970	Occupational Safety and Health Act
1970	Environmental Protection Agency
1971	Occupational Safety and Health Administration
1972	Consumer Product Safety Act
1972	Federal Insecticide, Fungicide, and Rodenticide Act
1972	Clean Water Act (1977 update)
1972	Noise Control Act
1973	Endangered Species Act
1974	Safe Drinking Water Act (1986, 1996 updates)
1975	Hazardous Materials Transportation Act
1976	Resource Conservation and Recovery Act
1976	Toxic Substances Control Act
1980	CERCLA (Superfund)
1984	Hazardous and Solid Waste Amendments
1986	Superfund Amendments and Reauthorization Act
1986	Emergency Wetlands Resources Act
1990	Oil Pollution Act
1993	North American Free Trade Agreement
2003	Healthy Forest Initiative

Attribution

Modified by Rachel Schleiger from Sustainability: A Comprehensive Foundation by Openstax (CC-BY).

This page titled 8.4: Environmental Policy is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **22.4: Environmental Policy** by Melissa Ha and Rachel Schleiger has no license indicated.

8.5: Sustainable Cities

Introduction

Sustainability, from science to philosophy to lifestyle, finds expression in the way we shape our cities. Cities are not just a collection of structures, but rather groups of people living different lifestyles together. When we ask if a lifestyle is sustainable, we're asking if it can endure. Some archaeologists posit that environmental imbalance doomed many failed ancient civilizations. What could the sustainable city look like, how would it function, and how can we avoid an imbalance that will lead to the collapse of our material civilization? While attempting to envision the "sustainable city" we must discern what factors will influence its shape and form in the future.

So what is the major hindrance for designing/updating for a "sustainable city"?

Sprawl: The outward expansion of human populations (typically in low-density residential and commercial development) into the outer edges of cities/towns.

As sprawl expands, it eats at natural and agricultural lands and isolates people. People have to travel further and further to go to work, the grocery store, a hospital, etcetera. What this means is that cities/towns have to find solutions to moving water, waste, power, services, road development/upkeep, etcetera to these communities. As such, there are chronic issues with sprawl on both sides that make opportunity and progress difficult due to economic constraints.

It's important to note that sprawl can happen at different scales:

- **Low-density sprawl:** Small 1-2 story buildings in high densities. Low-density sprawl can be further distinguished by income as neighborhoods tend to be segregated by this. (Ex: North America)
- **High-density sprawl:** High rise buildings in high densities (Ex: China)

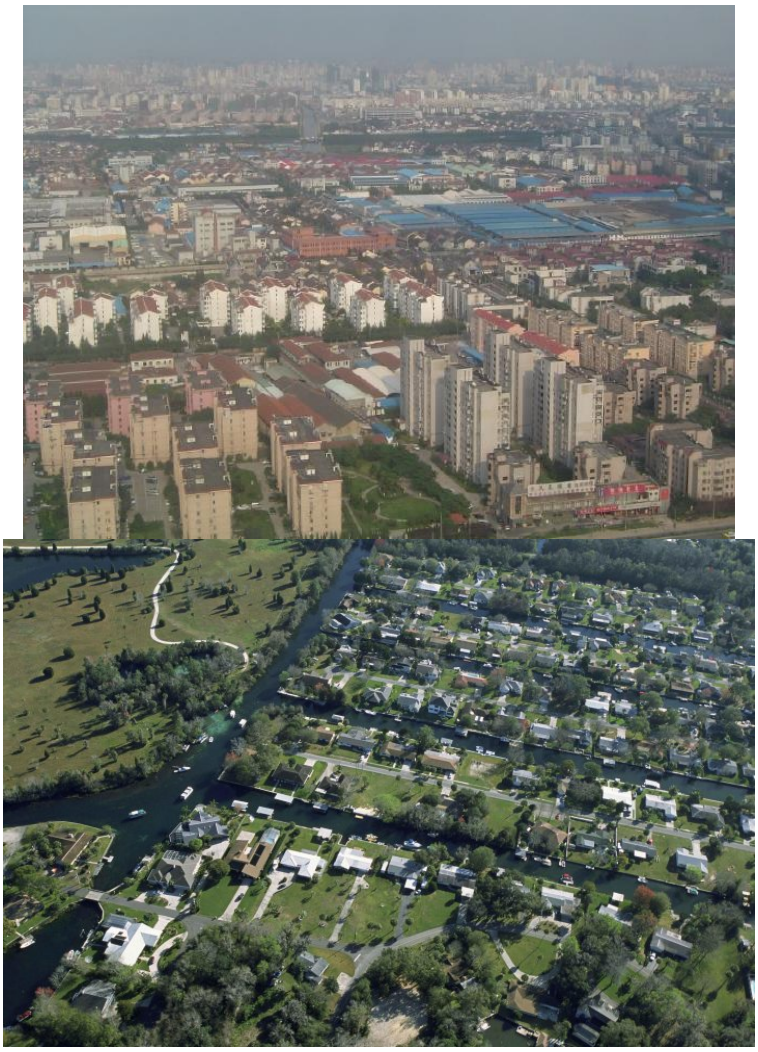


Figure 8.5.*a*: Left - High density sprawl in Shanghai China by John Patrick Robichaud in Wikimedia Commons (CC-BY), Right - Low density sprawl in North America by Pixnio (Public Domain).

So what can be done to design/update and move towards “sustainable cities”?

- **Mixed use development:** Lends residential, commercial, cultural, institutional, or industrial uses, where those functions are physically and functionally integrated, and that provides pedestrian connections. Making cities walkable!
- **Geographic equity:** Circulation system defined by ratios of how people move. This means more bikes, walking, public transportation (in many forms) and not cars. Equitability of the streets!
- **Preserve:** Preserve land, culture, and biodiversity as much as possible.
- **Mix**: Mix incomes, cultures, green spaces, public transport options, etcetera as much as possible. Allow people to connect and become neighbors not strangers.
- **Focus:** Match density of a community to meet transit/service needs.

Video 8.5.*a***:** Perspectives: Peter Calthrope describes 7 principles for building better cities.



Attribution

Rachel Schleiger modified from Environment and Society in a Changing World by PennState OER Initiative (CC-BY-NC-SA)

This page titled 8.5: Sustainable Cities is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **22.5: Sustainable Cities** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

8.6: Sustainability and the Future

Ecological Footprint

The **ecological footprint** (EF), developed by Canadian ecologist and planner William Rees, is basically an accounting tool that uses land as the unit of measurement to assess per capita consumption, production, and discharge needs. It starts from the assumption that every category of energy and material consumption and waste discharge requires the productive or absorptive capacity of a finite area of land or water. If we (add up) all the land requirements for all categories of consumption and waste discharge by a defined population, the total area represents the Ecological Footprint of that population on Earth whether or not this area coincides with the population's home region.

Land is used as the unit of measurement for the simple reason that, according to Rees, "Land area not only captures planet Earth's finiteness, it can also be seen as a proxy for numerous essential life support functions from gas exchange to nutrient recycling … land supports photosynthesis, the energy conduit for the web of life. Photosynthesis sustains all important food chains and maintains the structural integrity of ecosystems."

What does the ecological footprint tell us? Ecological footprint analysis can tell us in a vivid, ready-to-grasp manner how much of the Earth's environmental functions are needed to support human activities. It also makes visible the extent to which consumer lifestyles and behaviors are ecologically sustainable calculated that the ecological footprint of the average American is – conservatively – 5.1 hectares per capita of productive land. With roughly 7.4 billion hectares of the planet's total surface area of 51 billion hectares available for human consumption, if the current global population were to adopt American consumer lifestyles we would need two additional planets to produce the resources, absorb the wastes, and provide general life-support functions.

Common Cause and Sustainability

Common cause: Brings about change in all people/interest groups/world leaders because of common goals. Below are a list of sustainability goals across the sustainability paradigm. Most, if not all, are common cause goals for the now, and into the future.

Society

- No poverty
- Zero hunger
- Good health
- Quality education
- Gender equality
- Peace, justice, and strong institutions
- Clean water and sanitation
- Reduced inequalities

Environment

- Terrestrial life/habitats conservation and restoration
- Oceanic life/habitats conservation and restoration
- Climate action
- Reduced waste/pollution

Economy

- Affordable and clean energy
- Decent work and economic growth
- Industry, innovation, and infrastructure
- Responsible consumption and production
- Sustainable cities and communities
- Partnerships locally and globally

Sustainable Living

Sustainable living describes a lifestyle that attempts to reduce/eliminate an individuals (or societies) use of resources so as to be as close as possible to "net zero living". As such, an individual (or society) focuses on reducing their footprint (ecologically, carbon, socially, etc) through their choices/methods of:

- Use of resources (energy/diet/transportation/water/etc)
- Support of people/companies (voting/economic/social/etc)
- Practices of reducing, reusing, and recycling
- Consciousness of priority and focus for local concerns/needs versus global concerns/needs
- Sharing knowledge with their community (all ages)

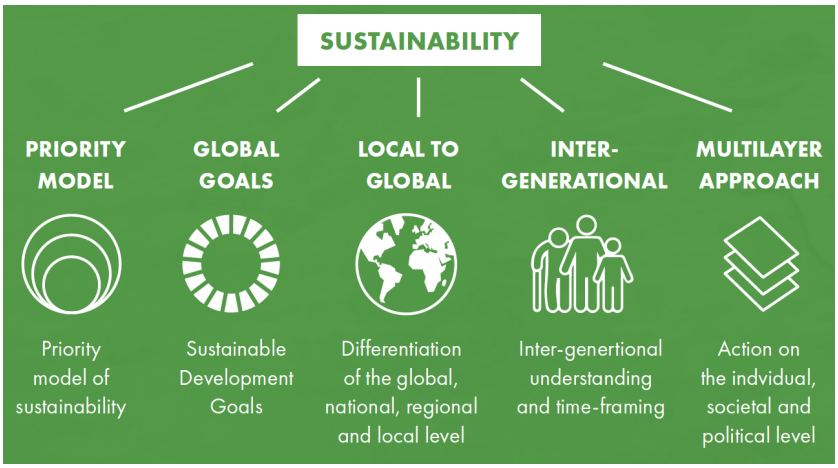


Figure 8.6.*a*: Considerations for sustainable living goals. Image by Sustainability Week Switzerland in Wikimedia Commons (CC-BY-SA)

The thought of getting started with sustainable living can be overwhelming! However, it is important to note that although we can acknowledge all the things that need to be accomplished, that that entire burden does not lie with just one individuals control (see figure below). It takes a multi-layerd approach to appropriately take action at the individual, societal, and political levels (see figure above).

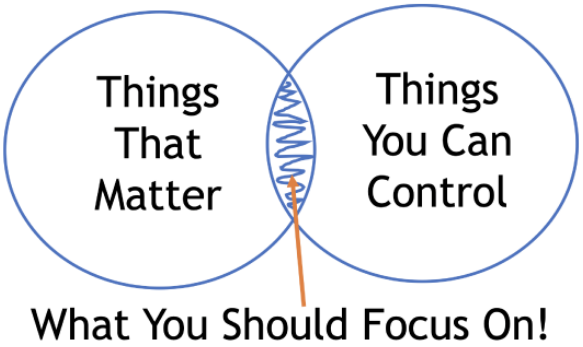


Figure 8.6.*b*: Focus for individual sustainability lifestyle goals. Image by Rachel Schleiger (CC-BY-NC)

The Bionic World

This book will conclude by stating a huge myth held by most of today's society, the myth of the **Bionic World**. It is the belief that science and technology will solve the pressing issues of human's impact on this earth. If society follows this logic, we have no tough choices to make about how we view and treat our surroundings, and decisions can be put off until the economic markets demand or justify a solution. Society can hope to be right…. However, until supporting evidence emerges that humanity can solve all their own problems with the ever shrinking timeline to do so, then we need to start making tough choices to ensure an equitable and sustainable future.



Figure 8.6.*c*: Robot arm holding up the Earth. Image by Pixabay (Public Domain)

Video 8.6.*a*: Perspectives: Michael Green discusses sustainability goals and what we can do moving forward (Videos from 2015 and 2018)





Attribution

Modified by Melissa Ha and Rachel Schleiger from Environment and Sustainability from *Environmental Biology* by Matthew R. Fisher (licensed under CC-BY) and The Myths of Restoration Ecology by Robert H. Hilderbrand, Adam C. Watts and April M. Randle 2005 (CC-BY-NC)

Tags recommended by the template: article:topic

This page titled 8.6: Sustainability and the Future is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **22.6: Sustainability and the Future** by Melissa Ha and Rachel Schleiger is licensed CC BY-NC 4.0.

8.7: Data Dive- Sustainable Investments

Overview

Climate change is the major focus of conversation across many different disciplines and interest groups. With a huge array of global sustainability challenges both now and in the future, investors are starting to re-evaluate where they want to spend their money. Today, sustainable investing is increasing exponentially! The graph below displays the global growth in sustainable investments between 2012 and 2018:

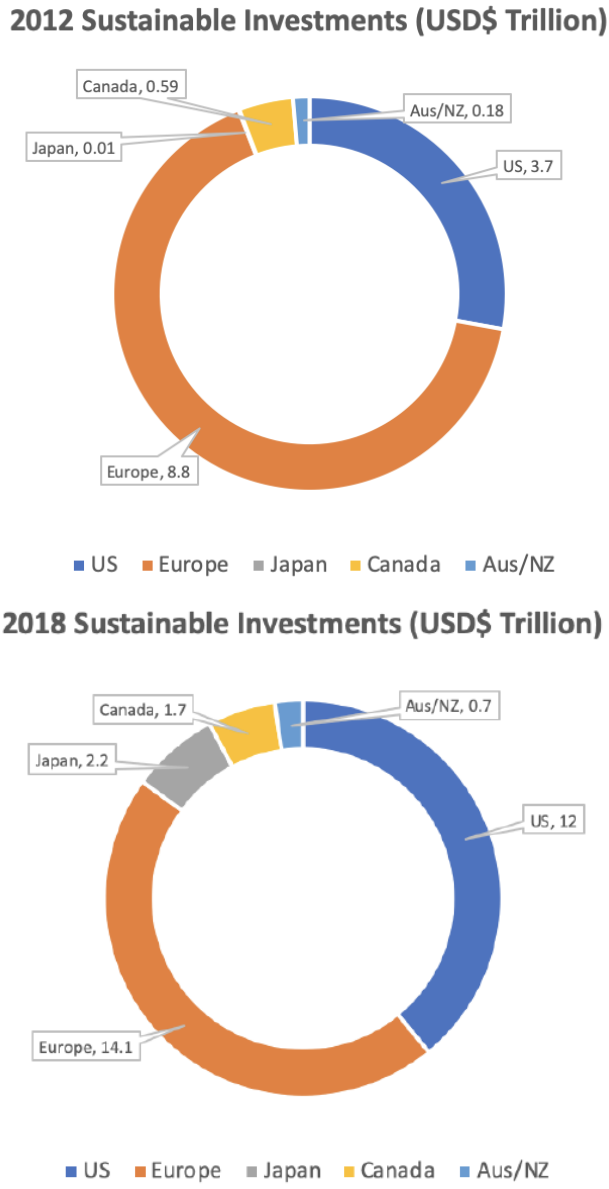


Figure 8.7.*a*: 2012 and 2018 global sustainable investments. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Global Sustainable Investment Alliance 2019.

Questions

1. What kind of graphs are these?
2. What major differences are there between the 2012 graph and the 2018 graph?
3. Which country(ies) is the major leader for sustainable investments?
4. Are there any countries/regions that are not represented on the graph that you would have predicted would be there?

5. Is there anything that surprises you about the results displayed in these graphs? Explain your answer.

Raw Data From Above Graph(s)

Table 8.7.*a***:** Raw data from 2012 and 2018 global sustainable investments. Graph by Rachel Schleiger (CC-BY-NC) modified from data in Global Sustainable Investment Alliance 2019.

Country/Continent	2012 Investments (USD$ Trillion)	2018 Investments (USD$ Trillion)
United States	3.7	12
Europe	8.8	14.1
Japan	0.01	2.2
Canada	0.59	1.7
Australia/New Zealand	0.18	0.7

Attributions

Rachel Schleiger (CC-BY-NC)

This page titled 8.7: Data Dive- Sustainable Investments is shared under a not declared license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **22.7: Data Dive- Sustainable Investments** by Melissa Ha and Rachel Schleiger has no license indicated.

8.8: Review

Summary

After completing this chapter you should be able to...

- Define sustainability and the sustainability paradigm
- Describe the historic events that lead to the sustainability paradigm
- Describe how ethics and moral extensionism connect to sustainability
- Understand that true sustainability promotes environmental equity and justice, thus eliminating environmental racism
- Describe the evolution of ethics leading to the modern environmental ethic
- Understand how environmental policy has changed based on movements and need in the US
- Define what sustainable cities are
- Describe how sprawl is a hindrance to sustainable cities
- Describe how sustainable lifestyle choices influence sustainability
- Acknowledge that the Bionic World is a myth

Sustainability is focused on meeting current needs without compromising the future generations to meet theirs. Sustainability is a paradigm created out of necessity from our past negligence to society, ecology, and economy (the three main components of sustainability).

Ethics plays a large role in addressing current and future sustainability goals necessitated from historic action/inaction. This is especially true for matters of **environmental justice and racism**, in addition to **environmental policies** and movements.

Sustainable cities are also discussed to address the large-scale sustainability concerns of how humanity lives. **Sprawl** is the biggest hindrance for creating the sustainable city. There are two main types of sprawl, **low-density** and **high-density**, that can be observed in many differently populated city layouts.

The **ecological footprint** is a tool to help illustrate the imbalance between lifestyles and behaviors and what is ecologically sustainable. In order to decrease ecological footprints, in addition to environmental injustices, sustainability promotes **common cause** sustainability goals. These emphasize that most goals humanity have are the same, and that these should be what both motivate and unite us as we move towards an uncertain future. In addition, humanity should also consider how we are approaching meeting these goals. The **myth of the bionic world** warns us not to make assumptions or take the easy way out. Only through tough decisions can we ensure an equitable and sustainable future.

Attribution

Rachel Schleiger (CC-BY-NC)

This page titled 8.8: Review is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Melissa Ha and Rachel Schleiger (ASCCC Open Educational Resources Initiative) .

- **22.8: Review** by Melissa Ha and Rachel Schleiger is licensed CC BY 4.0.

Index

A

acid rain
- 2.2.1: Acid Rain
- 2.2.1.2: Acid Rain Transport
- 2.2.1.3: Acid Snow
- 7.2.2.2: Acid Deposition

acid snow
- 2.2.1.3: Acid Snow

adaptive radiation
- 4.2.1: Extinction

Aeration
- 2.2.2.10: Water Treatment

age structure
- 5.1: The Human Population

agriculture
- 2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture

amphiprotic specie
- 2.2.2.8: Water Chemistry

anthrosphere
- 2.5.8.6: The Anthrosphere

Aposematic Coloration
- 3.3.1.2: Competition

aquatic biomes
- 3.5.3: Aquatic Biomes

Aquatic Chemistry
- 2.2.2: Aquatic Chemistry

aquifer
- 2.5.8.2: The Hydrosphere

atmosphere
- 2.2.3.1: Atmosphere

B

basic
- 2.2.4: Biochemical Cycles
- 2.2.4.2: Carbon Cycle
- 2.2.4.3: Nitrogen Cycle

Batesian mimicry
- 3.3.1.2: Competition

biochar
- 2.5.11.11: Soil Loss - Desertification and Deforestation

biochemical effect
- 2.5.16.4: Toxic Substances and Toxicology

biodegradation
- 2.5.12.10: Biodegradation

biodiversity
- 4.2.1: Extinction

biodiversity hotspot
- 4.2.1: Extinction

Biogeochemical Evolution
- 2.1.4.2: Biogeochemical Evolution

biological agents
- 5.4.1: Types of Environmental Hazards

biome
- 3.5.1: Climate and Biomes

biosphere
- 2.1.4: The Biosphere
- 2.5.8.5: The Biosphere
- 2.5.12.2: The Biosphere

biotransformation
- 2.6.6.12.1: Introduction to Biotransformation

bond energy
- 2.5.4.5: Covalent Bonds in Compounds

botulinum
- 2.5.16.6: Protecting Water, Food, and Air

breeder reactor
- 2.5.15.10: Nuclear Energy

C

camouflage
- 3.3.1.2: Competition

carbon cycle
- 2.2.3.2: Carbon Cycle
- 2.2.4.2: Carbon Cycle

carbon sequestration
- 2.5.10.5: The Enormous Importance of Climate

carrying capacity
- 3.2.3: Population Growth and Regulation

cellulose
- 2.5.14.7: Cellulose

Centers for Disease Control and Prevention
- 5.4.2: Epidemiology

CFCs
- 2.2.3.4: Ozone

chain reaction
- 2.5.15.10: Nuclear Energy

chemical diversity
- 4.2.1: Extinction

climate
- 2.5.10.5: The Enormous Importance of Climate

climate change
- 2.5.10.5: The Enormous Importance of Climate

climax community
- 3.3.1.2: Competition

clod
- 2.5.11.8: Have You Thanked a Clod Today?

coagulation
- 2.2.2.10: Water Treatment

cometabolism
- 2.5.12.10: Biodegradation

commensalism
- 3.3.1.2: Competition
- 3.3.1.4: Symbiosis

community
- 3.3.1: Biotic Interactions

community ecology
- 3.3.1: Biotic Interactions

competitive exclusion principle
- 3.3.1: Biotic Interactions
- 3.3.1.2: Competition

conservation tillage
- 2.5.11.11: Soil Loss - Desertification and Deforestation
- 2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture

Crop rotations
- 5.3.4: Sustainable Agriculture

Curtius rearrangement
- 2.4.3.4: Green Technologies for Safer Chemical Production

cyclodextrins
- 2.4.5.3: Specific Replacements

cytotoxicity assays
- 2.6.3.4: Different Cytotoxicity Assays

D

DDT
- 2.4.1.2: Green Chemistry Concepts
- 5.4.1: Types of Environmental Hazards

deforestation
- 2.5.11.11: Soil Loss - Desertification and Deforestation

desertification
- 2.5.11.11: Soil Loss - Desertification and Deforestation

detoxication
- 2.5.12.10: Biodegradation

DNA replication
- 2.6.2.3: DNA and RNA Metabolism

Dobson Unit (DU)
- 2.2.3.4: Ozone

E

ecological footprint
- 8.2: History of Sustainability

ecosystem diversity
- 4.2.1: Extinction

Emsleyan/Mertensian mimicry
- 3.3.1.2: Competition

Endemic diseases
- 5.4.2: Epidemiology

Endemic species
- 4.2.1: Extinction

Environmental Chemistry
- 2: Environmental Chemistry

environmental disturbance
- 3.3.1.2: Competition

Environmental Toxicology
- 5.4.4: Environmental Toxicology

Epidemic diseases
- 5.4.2: Epidemiology

epidemiology
- 5.4.2: Epidemiology

eutrophication
- 2.2.4.1: Biochemical Cycles

Exemplar
- 2.2.1.4: Electricity Generation
- 2.2.1.6: Sources of Nitrogen Oxides
- 2.2.1.7: Sources of Sulfur Oxides

exponential growth
- 3.2.3: Population Growth and Regulation

extinction
- 4.2.1: Extinction

extinction rate
- 4.2.1: Extinction

F

Flucculation
- 2.2.2.10: Water Treatment

food security
- 5.3.1: Food Security

foundation species
- 3.3.1.2: Competition

frontier ethic
8.3.2: Perspectives of the Environment

G

Gaia hypothesis
2.1.4.3: Gaia - Bioregulation of the Environment
genetic diversity
4.2.1: Extinction
genetic engineering
5.3.3: Selective Breeding and Genetic Engineering
geochemistry
2.1: Geochemistry (Lower)
geosphere
2.5.8.4: The Geosphere
2.5.11.2: The Nature of the Geosphere
glyphosate
2.5.12.1: Pigweed's Revenge
green chemistry
2.4: Key Elements of Green Chemistry (Lucia)
green products
2.5.1.10: Green Products and Services- Design for Sustainability

H

Herbivory
3.3.1: Biotic Interactions
heterogeneity
4.2.1: Extinction
host
3.3.1.2: Competition
human population
5.1: The Human Population
humification
2.5.11.8: Have You Thanked a Clod Today?
hydrogen
2.5.3.2: Hydrogen - The Simplest Atom
hydrologic cycle
2.5.8.2: The Hydrosphere
Hydrosphere
2.2.2.6: Natural Water

I

incidence
5.4.2: Epidemiology
integrated pest management
5.3.4: Sustainable Agriculture
intercropping
5.3.4: Sustainable Agriculture
island biogeography
3.3.1.2: Competition

K

keystone species
3.3.1.2: Competition
Koch's postulates
5.4.2: Epidemiology

L

Lattice Energy
2.5.4.3: Sodium Chloride and Ionic Bonds
life cycles
2.4.2.1: Introduction to Life Cycles
lignin
2.5.14.9: Lignin
logistic growth
3.2.3: Population Growth and Regulation
lysergic acid
2.4.1.2: Green Chemistry Concepts

M

Müllerian Mimicry
3.3.1.2: Competition
megafauna
4.2.1: Extinction
Mertensian mimicry
3.3.1.2: Competition
microcrystalline cellulose
2.5.14.7: Cellulose
micronutrients
2.5.11.9: Plant Nutrients and Fertilizers in Soil
mimicry
3.3.1: Biotic Interactions
mineralization
2.5.12.10: Biodegradation
minimal tillage
5.3.4: Sustainable Agriculture
moderator
2.5.15.10: Nuclear Energy
morbidity
5.4.2: Epidemiology
morbidity rates
5.4.2: Epidemiology
mortality
5.4.2: Epidemiology
mortality rates
5.4.2: Epidemiology
municipal refuse
2.5.11.7: The Geosphere as a Waste Repository
mutualism
3.3.1.2: Competition
3.3.1.4: Symbiosis

N

nuclear waste
2.5.15.10: Nuclear Energy
nucleosynthesis
2.1.1.2: Origin of the Elements

O

octet
2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons
octet rule
2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons
organic farming
5.3.4: Sustainable Agriculture
overpopulation
5.1.3: Looking Ahead
ozone
2.2.3.4: Ozone
ozone depletion
7.2.2.1: Ozone Depletion

P

paleosols
3.4.4: Soils
pandemic diseases
5.4.2: Epidemiology
parasite
3.3.1.2: Competition
parasitism
3.3.1.4: Symbiosis
passive stability
2.5.15.10: Nuclear Energy
Permanent Hard Water
2.2.2.6: Natural Water
photosynthesis
2.2.3.5: Photosynthesis
pic
2.5.7.11: Questions and Problems
pioneer species
3.3.1.2: Competition
population growth
3.2.3: Population Growth and Regulation
Prebiotic Atmosphere
2.1.3.2: Origin and Evolution of the Atmosphere
precautionary principle
8.2: History of Sustainability
predator
3.3.1: Biotic Interactions
prevalence
5.4.2: Epidemiology
primary succession
3.3.1.2: Competition
pseudogenes
3.3.1.4: Symbiosis

Q

QSAR
2.6.4.4: Other Approaches for Predictive Toxicity Modeling
Quantitative Structure Activity Relationships
2.6.4.4: Other Approaches for Predictive Toxicity Modeling

R

relative species abundance
3.3.1.2: Competition
Remediation
7.2.1.1: Water Pollutants and Their Sources
rhizosphere
2.5.11.8: Have You Thanked a Clod Today?
ricin
2.5.16.6: Protecting Water, Food, and Air

S

sanitary landfills
2.5.11.7: The Geosphere as a Waste Repository
sarin
2.5.16.6: Protecting Water, Food, and Air
secondary succession
3.3.1.2: Competition
sedimentation
2.2.2.10: Water Treatment

sevin
2.4.3.4: Green Technologies for Safer Chemical Production
soil conservation
2.5.11.11: Soil Loss - Desertification and Deforestation
soil erosion
2.5.11.11: Soil Loss - Desertification and Deforestation
Soil processes
3.4.4: Soils
Soil Profiles
3.4.4: Soils
species richness
3.3.1.2: Competition
Sporadic diseases
5.4.2: Epidemiology
sulfur oxides
2.2.1.7: Sources of Sulfur Oxides
surface albedo
2.5.11.2: The Nature of the Geosphere
sustainability
2.5.1.3: Seeing Green
sustainable agriculture
5.3.4: Sustainable Agriculture
symbiosis
3.3.1: Biotic Interactions
3.3.1.2: Competition
3.3.1.4: Symbiosis
T
Temporary Hard Water
2.2.2.6: Natural Water
terrestrial biomes
3.5.2: Terrestrial Biomes
Total Fertility Rate
5.1.3: Looking Ahead
transmutation
2.5.15.10: Nuclear Energy
travelling wave reactor
2.5.15.10: Nuclear Energy
W
water
2.2.2: Aquatic Chemistry
2.2.2.5: Fundamental Characteristics of Water
2.2.2.6: Natural Water
2.2.2.8: Water Chemistry
X
xenobiotic substances
2.5.16.4: Toxic Substances and Toxicology

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: ENVS C100: Environmental Science (Hoerer)

Webpages: 679

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC-SA 4.0: 38.7% (263 pages)
- CC BY-NC 4.0: 29.2% (198 pages)
- CC BY 4.0: 12.5% (85 pages)
- Undeclared: 11.5% (78 pages)
- CC BY-SA 4.0: 5.4% (37 pages)
- CC BY-NC 3.0: 2.7% (18 pages)

By Page

- ENVS C100: Environmental Science (Hoerer) - *Undeclared*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Introduction - *CC BY-NC 4.0*
    - 1.1: Environmental Science Preface - *Undeclared*
    - 1.2: What Is Science, and How Does It Work? - *CC BY-NC-SA 4.0*
      - 1.2.1: What is Science? - *CC BY-NC-SA 4.0*
      - 1.2.2: The Scientific Method - *CC BY-SA 4.0*
      - 1.2.3: Scientific Papers - *CC BY-NC 4.0*
      - 1.2.4: Basic and Applied Science - *CC BY-SA 4.0*
      - 1.2.5: Data Dive- Tonsillectomy Trends - *Undeclared*
      - 1.2.6: Review - *CC BY-NC 4.0*
    - 1.3: Math Blast- An Overview of Essential Mathematics Used in Science - *CC BY-NC 4.0*
      - 1.3.1: Getting down the basics- What are the essential mathematical and statistical terms and concepts? - *Undeclared*
      - 1.3.2: Summarizing the data- Descriptive statistics - *CC BY-NC 4.0*
      - 1.3.3: Putting it all together- Inferential statistics and hypothesis testing - *CC BY-NC 4.0*
      - 1.3.4: Data Dive- California Wildfires - *Undeclared*
      - 1.3.5: Review - *CC BY-NC 4.0*
  - 2: Environmental Chemistry - *CC BY-NC-SA 4.0*
- 00: Front Matter - *Undeclared*
  - TitlePage - *Undeclared*
  - InfoPage - *Undeclared*
  - Table of Contents - *Undeclared*
  - Licensing - *Undeclared*
- 2.1: Geochemistry (Lower) - *CC BY-NC 3.0*
  - Front Matter - *Undeclared*
    - Licensing - *Undeclared*
  - 2.1.1: The Earth and its Lithosphere - *CC BY-NC 3.0*
    - 2.1.1.1: Composition and Structure of the Earth - *CC BY-NC 3.0*
    - 2.1.1.2: Origin of the Elements - *CC BY-NC 3.0*
    - 2.1.1.3: Formation and evolution of the Earth - *CC BY-NC 3.0*
    - 2.1.1.4: The Earth's crust - *CC BY-NC 3.0*
  - 2.1.2: The Hydrosphere - *CC BY-NC 3.0*
    - 2.1.2.1: Water, Water Everywhere... - *CC BY-NC 3.0*
    - 2.1.2.2: The hydrosphere and the oceans - *CC BY-NC 3.0*
    - 2.1.2.3: Chemistry and geochemistry of the oceans - *CC BY-NC 3.0*
    - 2.1.2.4: Chemical budgets of oceanic elements - *CC BY-NC 3.0*
  - 2.1.3: The Atmosphere - *CC BY-NC 3.0*
    - 2.1.3.1: Structure and Composition of the Atmosphere - *CC BY-NC 3.0*

- 2.1.3.2: Origin and Evolution of the Atmosphere - *CC BY-NC 3.0*
- 2.1.4: The Biosphere - *CC BY-NC 3.0*
  - 2.1.4.1: Chemistry and Energetics of the Life Process - *CC BY-NC 3.0*
  - 2.1.4.2: Biogeochemical Evolution - *CC BY-NC 3.0*
  - 2.1.4.3: Gaia - Bioregulation of the Environment - *CC BY-NC 3.0*
- 2.2: Supplemental Modules (Environmental Chemistry) - *CC BY-NC-SA 4.0*
  - 2.2.1: Acid Rain - *CC BY-NC-SA 4.0*
    - 2.2.1.1: Acid Rain - *CC BY-NC-SA 4.0*
    - 2.2.1.2: Acid Rain Transport - *CC BY-NC-SA 4.0*
    - 2.2.1.3: Acid Snow - *CC BY-NC-SA 4.0*
    - 2.2.1.4: Electricity Generation - *CC BY-NC-SA 4.0*
    - 2.2.1.5: Playground 1 - *Undeclared*
    - 2.2.1.6: Sources of Nitrogen Oxides - *CC BY-NC-SA 4.0*
    - 2.2.1.7: Sources of Sulfur Oxides - *CC BY-NC-SA 4.0*
  - 2.2.2: Aquatic Chemistry - *CC BY-NC-SA 4.0*
    - 2.2.2.1: 1. Acid-Base Chemistry of Natural Aquatic Systems - *CC BY-NC-SA 4.0*
    - 2.2.2.2: 2. Carbonate Equilibria in Natural Waters - *CC BY-NC-SA 4.0*
    - 2.2.2.3: 3. Redox Equilibria in Natural Waters - *CC BY-NC-SA 4.0*
    - 2.2.2.4: 4. Solids in Contact With Natural Waters - *CC BY-NC-SA 4.0*
    - 2.2.2.5: Fundamental Characteristics of Water - *CC BY-NC-SA 4.0*
    - 2.2.2.6: Natural Water - *CC BY-NC-SA 4.0*
    - 2.2.2.7: Water Biology - *CC BY-NC-SA 4.0*
    - 2.2.2.8: Water Chemistry - *CC BY-NC-SA 4.0*
    - 2.2.2.9: Water Physics - *CC BY-NC-SA 4.0*
    - 2.2.2.10: Water Treatment - *CC BY-NC-SA 4.0*
  - 2.2.3: Atmospheric Chemistry - *CC BY-NC-SA 4.0*
    - 2.2.3.1: Atmosphere - *CC BY-NC-SA 4.0*
    - 2.2.3.2: Carbon Cycle - *CC BY-NC-SA 4.0*
    - 2.2.3.3: Carbon Oxides - *CC BY-NC-SA 4.0*
    - 2.2.3.4: Ozone - *CC BY-NC-SA 4.0*
    - 2.2.3.5: Photosynthesis - *CC BY-NC-SA 4.0*
  - 2.2.4: Biochemical Cycles - *CC BY-NC-SA 4.0*
    - 2.2.4.1: Biochemical Cycles - *CC BY-NC-SA 4.0*
    - 2.2.4.2: Carbon Cycle - *CC BY-NC-SA 4.0*
    - 2.2.4.3: Nitrogen Cycle - *CC BY-NC-SA 4.0*
- 2.3: Environmental Toxicology (van Gestel et al.) - *CC BY 4.0*
  - 00: Front Matter - *Undeclared*
    - Table of Contents - *Undeclared*
  - 2.3.1: Environmental Toxicology - *CC BY 4.0*
    - 2.3.1.1: Environmental toxicology - *CC BY 4.0*
    - 2.3.1.2: DPSIR - *CC BY 4.0*
    - 2.3.1.3: Short history - *CC BY 4.0*
  - 2.3.2: Environmental Chemistry, Chemicals - *CC BY 4.0*
    - 2.3.2.1: Introduction - *CC BY 4.0*
    - 2.3.2.2: Pollutants with specific properties - *CC BY 4.0*
    - 2.3.2.3: Pollutants with specific use - *CC BY 4.0*
  - 2.3.3: Environmental Chemistry - From Fate to Exposure - *CC BY 4.0*
    - 2.3.3.1: Environmental compartments - *CC BY 4.0*
    - 2.3.3.2: Sources of chemicals - *CC BY 4.0*
    - 2.3.3.3: Pathways and processes determining chemical fate - *CC BY 4.0*
    - 2.3.3.4: Partitioning and Partitioning Constants - *CC BY 4.0*
    - 2.3.3.5: Metal Speciation - *CC BY 4.0*
  - 2.3.4: Toxicology - *CC BY 4.0*
    - 2.3.4.1: Toxicokinetics - *CC BY 4.0*
    - 2.3.4.2: Toxicodynamics and Molecular Interactions - *CC BY 4.0*
    - 2.3.4.3: Toxicity Testing - *CC BY 4.0*
    - 2.3.4.4: Increasing ecological realism in toxicity testing - *CC BY 4.0*
    - 2.3.4.5: Availability and bioavailability - *CC BY 4.0*
    - 2.3.4.6: Degradation - *CC BY 4.0*
    - 2.3.4.7: Modelling exposure - *CC BY 4.0*
    - 2.3.4.8: Section 4.1. Toxicokinetics - *CC BY 4.0*
    - 2.3.4.9: Section 4.4. Increasing ecological realism in toxicity testing - *CC BY 4.0*
    - 2.3.4.10: Section 4.2. Toxicodynamics and Molecular Interactions - *CC BY 4.0*
    - 2.3.4.11: Section 4.3. Toxicity testing - *CC BY 4.0*
  - 2.3.5: Population, Community and Ecosystem Ecotoxicology - *CC BY 4.0*

- 2.3.5.1: Introduction- Linking population, community and ecosystem responses - *CC BY 4.0*
- 2.3.5.2: Population ecotoxicology in laboratory settings - *CC BY 4.0*
- 2.3.5.3: Wildlife population ecotoxicology - *CC BY 4.0*
- 2.3.5.4: Trait-based approaches - *CC BY 4.0*
- 2.3.5.5: Population models - *CC BY 4.0*
- 2.3.5.6: Metapopulations - *CC BY 4.0*
- 2.3.5.7: Community ecotoxicology - *CC BY 4.0*
- 2.3.5.8: Structure versus function incl. ecosystem services - *CC BY 4.0*
- 2.3.5.9: Landscape ecotoxicology - *CC BY 4.0*

- 2.3.6: Risk Assessment and Regulation - *CC BY 4.0*
  - 2.3.6.1: Introduction - The Essence of Risk Assessment - *CC BY 4.0*
  - 2.3.6.2: Ecosystem services and protection goals - *CC BY 4.0*
  - 2.3.6.3: Predictive risk assessment approaches and tools - *CC BY 4.0*
  - 2.3.6.4: Diagnostic risk assessment approaches and tools - *CC BY 4.0*
  - 2.3.6.5: Regulatory Frameworks - *CC BY 4.0*
  - 2.3.6.6: Risk management and risk communication - *CC BY 4.0*
  - 2.3.6.7: Risk perception - *CC BY 4.0*

- 2.4: Key Elements of Green Chemistry (Lucia) - *CC BY 4.0*
  - Front Matter - *CC BY-NC-SA 4.0*
    - Licensing - *Undeclared*
  - 2.4.1: Key Elements of Green Chemistry - *CC BY 4.0*
    - 2.4.1.1: Introduction to Green Chemistry - *CC BY 4.0*
    - 2.4.1.2: Green Chemistry Concepts - *CC BY 4.0*
    - 2.4.1.3: Case Study - *CC BY 4.0*
    - 2.4.1.4: Conclusions and Review - *CC BY 4.0*
  - 2.4.2: Life-Cycle Analysis - *CC BY 4.0*
    - 2.4.2.1: Introduction to Life Cycles - *CC BY 4.0*
    - 2.4.2.2: LCA/LCIA Concepts - *CC BY 4.0*
    - 2.4.2.3: Conclusions and Review - *CC BY 4.0*
  - 2.4.3: Hazards - *CC BY 4.0*
    - 2.4.3.1: Introduction to Hazards of Chemistry - *CC BY 4.0*
    - 2.4.3.2: Hazard Concepts - *CC BY 4.0*
    - 2.4.3.3: Case Study - Badger Army Ammunition Plant - *CC BY 4.0*
    - 2.4.3.4: Green Technologies for Safer Chemical Production - *CC BY 4.0*
    - 2.4.3.5: Conclusions and Review Questions - *CC BY 4.0*
  - 2.4.4: Alternative Solvents - *CC BY 4.0*
    - 2.4.4.1: Introduction to Solvents - *CC BY 4.0*
    - 2.4.4.2: Solvent Concepts - *CC BY 4.0*
    - 2.4.4.3: Solvent Alternatives - *CC BY 4.0*
    - 2.4.4.4: Conclusions and Review - *CC BY 4.0*
  - 2.4.5: Alternative Reagants - *CC BY 4.0*
    - 2.4.5.1: Introduction to Reagants - *CC BY 4.0*
    - 2.4.5.2: Reagant Concepts - *CC BY 4.0*
    - 2.4.5.3: Specific Replacements - *CC BY 4.0*
    - 2.4.5.4: Conclusions and Review - *CC BY 4.0*
  - 2.4.6: Reaction Types, Design, and Efficiency - *CC BY 4.0*
    - 2.4.6.1: Introduction to Reactions - *CC BY 4.0*
    - 2.4.6.2: Reaction Concepts - *CC BY 4.0*
    - 2.4.6.3: Reaction Design Concepts - *CC BY 4.0*
    - 2.4.6.4: Conclusions and Review - *CC BY 4.0*
  - Back Matter - *CC BY-NC-SA 4.0*
    - Detailed Licensing - *Undeclared*

- 2.5: Green Chemistry and the Ten Commandments of Sustainability (Manahan) - *CC BY-NC-SA 4.0*
  - Front Matter - *CC BY-NC-SA 4.0*
    - Licensing - *Undeclared*
  - 2.5.1: Sustainability and the Environment - *CC BY-NC-SA 4.0*
    - 2.5.1.1: Sustainability - *CC BY-NC-SA 4.0*
    - 2.5.1.2: The Environment and the Five Environmental Sphere - *CC BY-NC-SA 4.0*
    - 2.5.1.3: Seeing Green - *CC BY-NC-SA 4.0*
    - 2.5.1.4: Natural Capital of the Earth - *CC BY-NC-SA 4.0*
    - 2.5.1.5: Sustainability as a Group Effort- It Takes a (Very Big) Village - *CC BY-NC-SA 4.0*
    - 2.5.1.6: Sustainable Energy- Away from the Sun and Back Again - *CC BY-NC-SA 4.0*
    - 2.5.1.7: Green Science - *CC BY-NC-SA 4.0*
    - 2.5.1.8: Green Technology - *CC BY-NC-SA 4.0*
    - 2.5.1.9: Sustainability and the Eco-Economy - *CC BY-NC-SA 4.0*

- 2.5.1.10: Green Products and Services- Design for Sustainability - *CC BY-NC-SA 4.0*
- 2.5.1.11: Questions and Problems - *CC BY-NC-SA 4.0*

- 2.5.2: The Key Role of Chemistry and Making Chemistry Green - *CC BY-NC-SA 4.0*
  - 2.5.2.1: Chemistry is Good (and Unavoidable) - *CC BY-NC-SA 4.0*
  - 2.5.2.2: The Environment and its Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.2.3: What is Environmental Chemistry? - *CC BY-NC-SA 4.0*
  - 2.5.2.4: Environmental Pollution - *CC BY-NC-SA 4.0*
  - 2.5.2.5: Practice of Green Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.2.6: Green Chemistry and Synthetic Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.2.7: Reduction of Risk- Hazard and Exposure - *CC BY-NC-SA 4.0*
  - 2.5.2.8: The Risks of No Risks - *CC BY-NC-SA 4.0*
  - 2.5.2.9: Waste Prevention - *CC BY-NC-SA 4.0*
  - 2.5.2.10: Basic Principles of Green Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.2.11: Some Things to Know About Chemistry before You Even Start - *CC BY-NC-SA 4.0*
  - 2.5.2.12: Combining Atoms to Make Molecules and Compounds - *CC BY-NC-SA 4.0*
  - 2.5.2.13: The process of making and breaking chemical bonds- chemical reactions - *CC BY-NC-SA 4.0*
  - 2.5.2.14: The Nature of Matter and States of Matter - *CC BY-NC-SA 4.0*
  - 2.5.2.15: References and Questions - *CC BY-NC-SA 4.0*
- 2.5.3: The Elements - Basic Building Blocks of Green Chemicals - *CC BY-NC-SA 4.0*
  - 2.5.3.1: Elements, Atoms, and Atomic Theory - *CC BY-NC-SA 4.0*
  - 2.5.3.2: Hydrogen - The Simplest Atom - *CC BY-NC-SA 4.0*
  - 2.5.3.3: Helium - The First Noble Gas - *CC BY-NC-SA 4.0*
  - 2.5.3.4: Lithium, The First Metal - *CC BY-NC-SA 4.0*
  - 2.5.3.5: The Second Period of the Periodic Table - *CC BY-NC-SA 4.0*
  - 2.5.3.6: The Magic Octet of 8 Outer-Shell Electrons - *CC BY-NC-SA 4.0*
  - 2.5.3.7: Completing the 20-Element Periodic Table - *CC BY-NC-SA 4.0*
  - 2.5.3.8: The Brief Periodic Table is Complete - *CC BY-NC-SA 4.0*
  - 2.5.3.E: The Elements - Basic Building Blocks of Green Chemicals (Exercises) - *CC BY-NC-SA 4.0*
  - 2.5.3.R: References - *CC BY-NC-SA 4.0*
- 2.5.4: Compounds- Safer Materials for a Safer World - *CC BY-NC-SA 4.0*
  - 2.5.4.1: Chemical Bonds and Compound Formation - *CC BY-NC-SA 4.0*
  - 2.5.4.2: Electrons Involved in Chemical Bonds and Octets of Electrons - *CC BY-NC-SA 4.0*
  - 2.5.4.3: Sodium Chloride and Ionic Bonds - *CC BY-NC-SA 4.0*
  - 2.5.4.4: Covalent Bonds in H2 and Other Molecules - *CC BY-NC-SA 4.0*
  - 2.5.4.5: Covalent Bonds in Compounds - *CC BY-NC-SA 4.0*
  - 2.5.4.6: Covalent Bonds and Green Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.4.7: Predicting Formulas of Covalently Bound Compounds - *CC BY-NC-SA 4.0*
  - 2.5.4.8: Chemical Formulas, the Mole, and Percentage Composition - *CC BY-NC-SA 4.0*
  - 2.5.4.9: What Are Chemical Compounds Called? - *CC BY-NC-SA 4.0*
  - 2.5.4.10: Acids, Bases and Salts - *CC BY-NC-SA 4.0*
  - 2.5.4.11: Questions and Problems - *CC BY-NC-SA 4.0*
- 2.5.5: Chemical Reactions- Making Materials Safely and Sustainable - *CC BY-NC-SA 4.0*
  - 2.5.5.1: Describing What Happens with Chemical Equations - *CC BY-NC-SA 4.0*
  - 2.5.5.2: Balancing Chemical Equations - *CC BY-NC-SA 4.0*
  - 2.5.5.3: Just Because You Can Write It Doesn't Mean That It Will Happen - *CC BY-NC-SA 4.0*
  - 2.5.5.4: Yield And Atom Economy in Chemical Reactions - *CC BY-NC-SA 4.0*
  - 2.5.5.5: Catalysts That Make Reactions Go - *CC BY-NC-SA 4.0*

- 2.5.5.6: Kinds of Chemical Reactions - *CC BY-NC-SA 4.0*
- 2.5.5.7: Oxidation-Reduction Reactions and Green Chemistry - *CC BY-NC-SA 4.0*
- 2.5.5.8: Quantitative Information from Chemical Reactions - *CC BY-NC-SA 4.0*
- 2.5.5.9: Energy in Chemical Reactions - *CC BY-NC-SA 4.0*
- 2.5.5.10: Stoichiometry by the Mole Ratio Method - *CC BY-NC-SA 4.0*
- 2.5.5.11: Limiting Reactant and Percent Yield - *CC BY-NC-SA 4.0*
- 2.5.5.12: Titrations - Measuring Moles by Volume of Solution - *CC BY-NC-SA 4.0*
- 2.5.5.13: Industrial Chemical Reactions - The Solvay Process - *CC BY-NC-SA 4.0*
- 2.5.5.14: Questions and Problems - *CC BY-NC-SA 4.0*

- 2.5.6: The Wonderful World of Carbon - Organic Chemistry and Biochemicals - *CC BY-NC-SA 4.0*
  - 2.5.6.1: Rings and Chains of Carbon Atoms - *CC BY-NC-SA 4.0*
  - 2.5.6.2: Compounds of Carbon and Hydrogen - Hydrocarbons - *CC BY-NC-SA 4.0*
  - 2.5.6.3: Using Lines To Show Organic Structural Formulas - *CC BY-NC-SA 4.0*
  - 2.5.6.4: Functional Groups - *CC BY-NC-SA 4.0*
  - 2.5.6.5: Giant Molecules from Small Organic Molecules - *CC BY-NC-SA 4.0*
  - 2.5.6.6: Questions and Problems - *CC BY-NC-SA 4.0*
  - 2.5.6.7: Supplementary References - *CC BY-NC-SA 4.0*
- 2.5.7: Chemistry of Life and Green Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.7.1: Green Biochemistry - *CC BY-NC-SA 4.0*
  - 2.5.7.2: Biochemistry and the Cells - *CC BY-NC-SA 4.0*
  - 2.5.7.3: Carbohydrates - *CC BY-NC-SA 4.0*
  - 2.5.7.4: Proteins - *CC BY-NC-SA 4.0*
  - 2.5.7.5: Lipids - Fats, Oils, and Hormones - *CC BY-NC-SA 4.0*
  - 2.5.7.6: Nucleic Acids - *CC BY-NC-SA 4.0*
  - 2.5.7.7: Enzymes - *CC BY-NC-SA 4.0*
  - 2.5.7.8: Biochemical Processes in Metabolism - *CC BY-NC-SA 4.0*
  - 2.5.7.9: Biochemistry of Toxic Substances and Toxicological Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.7.10: Literature Cited - *CC BY-NC-SA 4.0*
  - 2.5.7.11: Questions and Problems - *CC BY-NC-SA 4.0*
  - 2.5.7.12: Supplementary References - *CC BY-NC-SA 4.0*
- 2.5.8: The Five Environmental Spheres and Biogeochemical Cycle - *CC BY-NC-SA 4.0*
  - 2.5.8.1: Five Environmental Spheres - *CC BY-NC-SA 4.0*
  - 2.5.8.2: The Hydrosphere - *CC BY-NC-SA 4.0*
  - 2.5.8.3: The Atmosphere - *CC BY-NC-SA 4.0*
  - 2.5.8.4: The Geosphere - *CC BY-NC-SA 4.0*
  - 2.5.8.5: The Biosphere - *CC BY-NC-SA 4.0*
  - 2.5.8.6: The Anthrosphere - *CC BY-NC-SA 4.0*
  - 2.5.8.7: Cycles of Matter - *CC BY-NC-SA 4.0*
  - 2.5.8.8: Questions and Problems - *CC BY-NC-SA 4.0*
  - 2.5.8.9: Supplementary References - *CC BY-NC-SA 4.0*
- 2.5.9: Water - the Ultimate Green Substance - *CC BY-NC-SA 4.0*
  - 2.5.9.1: H_2O - Simple Formula, Remarkable Molecule - *CC BY-NC-SA 4.0*
  - 2.5.9.2: Occurrence, Availability , and Utilization of Water on Earth - *CC BY-NC-SA 4.0*
  - 2.5.9.3: Water Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.9.4: Water Pollution - *CC BY-NC-SA 4.0*
  - 2.5.9.5: Greening of Water - Purification Before Use - *CC BY-NC-SA 4.0*
  - 2.5.9.6: Wastewater Treatment - *CC BY-NC-SA 4.0*
  - 2.5.9.7: Advanced Wastewater Treatment and Recycling - *CC BY-NC-SA 4.0*
  - 2.5.9.8: The Many Uses of Water - *CC BY-NC-SA 4.0*
  - 2.5.9.9: Hot Water - Pressurized Subcritical Water - *CC BY-NC-SA 4.0*
  - 2.5.9.10: Supercritical Water - *CC BY-NC-SA 4.0*
  - 2.5.9.11: Literature Cited - *CC BY-NC-SA 4.0*
  - 2.5.9.12: Questions and Problems - *CC BY-NC-SA 4.0*
  - 2.5.9.13: Supplementary References - *CC BY-NC-SA 4.0*
- 2.5.10: Blue Skies for a Green Environment - *CC BY-NC-SA 4.0*
  - 2.5.10.1: More Than Just Air to Breathe - *CC BY-NC-SA 4.0*

- 2.5.10.2: The Gas Laws - *CC BY-NC-SA 4.0*
- 2.5.10.3: The Protective Atmosphere - *CC BY-NC-SA 4.0*
- 2.5.10.4: Atmospheric Chemistry and Photochemical Reactions - *CC BY-NC-SA 4.0*
- 2.5.10.5: The Enormous Importance of Climate - *CC BY-NC-SA 4.0*
- 2.5.10.6: Atmospheric Particle Pollutants - *CC BY-NC-SA 4.0*
- 2.5.10.7: Pollutant Gaseous Oxides - *CC BY-NC-SA 4.0*
- 2.5.10.8: Acid Rain - *CC BY-NC-SA 4.0*
- 2.5.10.9: Miscellaneous Gases in the Atmosphere - *CC BY-NC-SA 4.0*
- 2.5.10.10: Photochemical Smog - *CC BY-NC-SA 4.0*
- 2.5.10.11: Natural Capital of the Atmosphere - *CC BY-NC-SA 4.0*
- 2.5.10.12: Energy and Mass Transfer in the Atmosphere - *CC BY-NC-SA 4.0*
- 2.5.10.13: Questions and Problems - *Undeclared*
- 2.5.10.14: Supplementary References - *Undeclared*

- 2.5.11: The Geosphere and a Green Earth - *CC BY-NC-SA 4.0*
  - 2.5.11.1: Probing Deep into the Geosphere - A Well Too Far? - *CC BY-NC-SA 4.0*
  - 2.5.11.2: The Nature of the Geosphere - *CC BY-NC-SA 4.0*
  - 2.5.11.3: The Geosphere as a Source of Natural Capital - *CC BY-NC-SA 4.0*
  - 2.5.11.4: Environmental Hazards of the Geosphere - *CC BY-NC-SA 4.0*
  - 2.5.11.5: Water on and in the Geosphere - *CC BY-NC-SA 4.0*
  - 2.5.11.6: Anthrospheric Influences on the Geosphere - *CC BY-NC-SA 4.0*
  - 2.5.11.7: The Geosphere as a Waste Repository - *CC BY-NC-SA 4.0*
  - 2.5.11.8: Have You Thanked a Clod Today? - *CC BY-NC-SA 4.0*
  - 2.5.11.9: Plant Nutrients and Fertilizers in Soil - *CC BY-NC-SA 4.0*
  - 2.5.11.10: Soil and Plants Related To Wastes and Pollutants - *CC BY-NC-SA 4.0*
  - 2.5.11.11: Soil Loss - Desertification and Deforestation - *CC BY-NC-SA 4.0*
  - 2.5.11.12: Literature Cited and Supplementary References - *Undeclared*
  - 2.5.11.13: Questions and Problems - *Undeclared*
- 2.5.12: The Biosphere and the Role of Green Chemistry in Feeding a Hungry World - *CC BY-NC-SA 4.0*
  - 2.5.12.1: Pigweed's Revenge - *CC BY-NC-SA 4.0*
  - 2.5.12.2: The Biosphere - *CC BY-NC-SA 4.0*
  - 2.5.12.3: Cells - Basic Units of Life - *CC BY-NC-SA 4.0*
  - 2.5.12.4: Metabolism and Control in Organisms - *CC BY-NC-SA 4.0*
  - 2.5.12.5: Reproduction and Inherited Traits - *CC BY-NC-SA 4.0*
  - 2.5.12.6: Stability and Equilibrium of the Biosphere - *CC BY-NC-SA 4.0*
  - 2.5.12.7: DNA and the Human Genome - *CC BY-NC-SA 4.0*
  - 2.5.12.8: Genetic Engineering - *CC BY-NC-SA 4.0*
  - 2.5.12.9: Biological Interaction with Environmental Chemicals - *CC BY-NC-SA 4.0*
  - 2.5.12.10: Biodegradation - *CC BY-NC-SA 4.0*
  - 2.5.12.11: Production of Food and Fiber by the Biosphere - Agriculture - *CC BY-NC-SA 4.0*
  - 2.5.12.12: Agricultural Applications of Genetically Modified Organisms - *CC BY-NC-SA 4.0*
  - 2.5.12.13: The Anthrosphere in Support of the Biosphere - *CC BY-NC-SA 4.0*
  - 2.5.12.14: Livestock and their Wastes - *CC BY-NC-SA 4.0*
  - 2.5.12.15: Questions and Problems - *Undeclared*
  - 2.5.12.16: Supplementary References - *Undeclared*
- 2.5.13: The Anthrosphere, Industrial Ecology, and Green Chemistry - *CC BY-NC-SA 4.0*
  - 2.5.13.1: Industrial Ecology and Industrial Ecosystems - *CC BY-NC-SA 4.0*
  - 2.5.13.2: Metabolic Processes in Industrial Ecosystems - *CC BY-NC-SA 4.0*
  - 2.5.13.3: Life Cycles in Industrial Ecosystems - *CC BY-NC-SA 4.0*
  - 2.5.13.4: Kinds of Products - *CC BY-NC-SA 4.0*

- 2.5.13.5: Attributes Required by an Industrial Ecosystem - *CC BY-NC-SA 4.0*
- 2.5.13.6: The Kalundborg Industrial Ecosystem - *CC BY-NC-SA 4.0*
- 2.5.13.7: Environmental Impacts of Industrial Ecosystems - *CC BY-NC-SA 4.0*
- 2.5.13.8: Green Chemistry and Industrial Ecology - *CC BY-NC-SA 4.0*
- 2.5.13.9: Predicting and Reducing Hazards - *CC BY-NC-SA 4.0*
- 2.5.13.10: The E-Factor in Green Chemistry - *CC BY-NC-SA 4.0*
- 2.5.13.11: Catalysts and Catalysis - *CC BY-NC-SA 4.0*
- 2.5.13.12: Biocatalysis with Enzymes - *CC BY-NC-SA 4.0*
- 2.5.13.13: Energizing Chemical Reactions and Process Intensification - *CC BY-NC-SA 4.0*
- 2.5.13.14: Solvents and Alternate Reaction Media - *CC BY-NC-SA 4.0*
- 2.5.13.15: Feedstocks and Reagents - *CC BY-NC-SA 4.0*
- 2.5.13.16: Literature Cited and Supplementary References - *Undeclared*
- 2.5.13.17: Questions and Problems - *Undeclared*

- 2.5.14: Feeding the Anthrosphere- Utilizing Renewable and Biological Materials - *CC BY-NC-SA 4.0*
  - 2.5.14.1: Feeding the Anthrosphere - *CC BY-NC-SA 4.0*
  - 2.5.14.2: Utilization of Feedstocks - *CC BY-NC-SA 4.0*
  - 2.5.14.3: Biological Feedstocks - *CC BY-NC-SA 4.0*
  - 2.5.14.4: Biological Sources of Chemicals - *CC BY-NC-SA 4.0*
  - 2.5.14.5: Biorefineries and Biomass Utilization - *CC BY-NC-SA 4.0*
  - 2.5.14.6: Monosaccharide Feedstocks - Glucose, Fructose, and Xylose - *CC BY-NC-SA 4.0*
  - 2.5.14.7: Cellulose - *CC BY-NC-SA 4.0*
  - 2.5.14.8: Feedstocks from Cellulose Wastes - *CC BY-NC-SA 4.0*
  - 2.5.14.9: Lignin - *CC BY-NC-SA 4.0*
  - 2.5.14.10: Direct Biosynthesis of Polymers - *CC BY-NC-SA 4.0*
  - 2.5.14.11: Bioconversion Processes for Synthetic Chemicals - *Undeclared*
  - 2.5.14.12: Bamboo - Ancient Material for the Future - *CC BY-NC-SA 4.0*
  - 2.5.14.13: Literature Cited - *Undeclared*
  - 2.5.14.14: Questions and Problems - *Undeclared*
- 2.5.15: Sustainable Energy- The Essential Basis of Green Systems - *CC BY-NC-SA 4.0*
  - 2.5.15.1: Sustainability and Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.2: What is Energy? - *CC BY-NC-SA 4.0*
  - 2.5.15.3: Radiant Energy from the Sun - *CC BY-NC-SA 4.0*
  - 2.5.15.4: Sources of Energy Used in the Anthrosphere - *CC BY-NC-SA 4.0*
  - 2.5.15.5: Conversions Between Forms of Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.6: Green Technology for Energy Conversion - *CC BY-NC-SA 4.0*
  - 2.5.15.7: Energy Conservation - *CC BY-NC-SA 4.0*
  - 2.5.15.8: Depletable Fossil Fuels - *CC BY-NC-SA 4.0*
  - 2.5.15.9: Carbon Sequestration for Fossil Fuel Utilization - *CC BY-NC-SA 4.0*
  - 2.5.15.10: Nuclear Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.11: Renewable Energy Sources - Solar Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.12: Energy from Wind and Water - *CC BY-NC-SA 4.0*
  - 2.5.15.13: Biomass Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.14: Geothermal Energy - *CC BY-NC-SA 4.0*
  - 2.5.15.15: Hydrogen for Energy Storage and Utilization - *CC BY-NC-SA 4.0*
  - 2.5.15.16: Literature Cited and Supplementary References - *Undeclared*
  - 2.5.15.17: Questions and Problems - *Undeclared*
- 2.5.16: Terrorism, Toxicity, and Vulnerability- Green Chemistry and Technology in Defense of Human Welfare - *CC BY-NC-SA 4.0*
  - 2.5.16.1: Vulnerability to Terrorist Attack - *CC BY-NC-SA 4.0*
  - 2.5.16.2: Protecting the Anthrosphere - *CC BY-NC-SA 4.0*

- 2.5.16.3: Substances That Explode, Burn, or React Violently - *CC BY-NC-SA 4.0*
- 2.5.16.4: Toxic Substances and Toxicology - *CC BY-NC-SA 4.0*
- 2.5.16.5: Toxic Chemical Attack - *CC BY-NC-SA 4.0*
- 2.5.16.6: Protecting Water, Food, and Air - *CC BY-NC-SA 4.0*
- 2.5.16.7: Detecting Hazards - *CC BY-NC-SA 4.0*
- 2.5.16.8: Green Chemistry to Combat Terrorism - *CC BY-NC-SA 4.0*
- 2.5.16.9: Green Chemistry for Sustainable Prosperity and a Safer World - *CC BY-NC-SA 4.0*
- 2.5.16.10: Green Chemistry to Combat Terrorism - *Undeclared*
- 2.5.16.11: Green Chemistry for Sustainable Prosperity and and a Safer World - *Undeclared*
- 2.5.16.12: Literature Cited - *Undeclared*
- 2.5.16.13: Questions and Problems - *Undeclared*

- 2.5.17: The Ten Commandments of Sustainability and Sensible Measure - *CC BY-NC-SA 4.0*
  - 2.5.17.1: We Cannot Go on Like This - *CC BY-NC-SA 4.0*
  - 2.5.17.2: The First Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.3: The Second Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.4: The Third Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.5: The Fourth Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.6: The Fifth Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.7: The Sixth Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.8: The Seventh Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.9: The Ninth Commandment - *Undeclared*
  - 2.5.17.10: The Tenth Commandment - *Undeclared*
  - 2.5.17.11: Some Sensible Measures for Sustainability - *Undeclared*
  - 2.5.17.12: The Eighth Commandment - *CC BY-NC-SA 4.0*
  - 2.5.17.13: Literature Cited - *Undeclared*
  - 2.5.17.14: Questions and Problems - *Undeclared*
  - 2.5.17.15: Supplementary References - *Undeclared*
- Back Matter - *CC BY-NC-SA 4.0*
  - Detailed Licensing - *Undeclared*

- 2.6: Toxicology MSDT - *CC BY-NC 4.0*
  - Front Matter - *CC BY-NC 4.0*
    - Licensing - *Undeclared*
    - Preface - *CC BY-NC 4.0*
  - 2.6.1: Pathophysiology - *CC BY-NC 4.0*
    - 2.6.1.1: What is Pathophysiology? - *CC BY-NC 4.0*
    - 2.6.1.2: Outcomes of Targeted and Non-Targeted Toxicity - *CC BY-NC 4.0*
    - 2.6.1.3: Cellular Response to Toxicant-Induced Injury - *CC BY-NC 4.0*
    - 2.6.1.4: Repair and Adaptation - *CC BY-NC 4.0*
    - 2.6.1.5: Patterns of Toxic Injury - *CC BY-NC 4.0*
    - 2.6.1.6: Final Evaluation - *CC BY-NC 4.0*
  - 2.6.2: Biochemistry and Molecular Genetics - *CC BY-NC 4.0*
    - 2.6.2.1: Introduction to Biomolecules and Cell Components - *CC BY-NC 4.0*
    - 2.6.2.2: Cell Structure and Subcellular Compartments - *CC BY-NC 4.0*
    - 2.6.2.3: DNA and RNA Metabolism - *CC BY-NC 4.0*
    - 2.6.2.4: Epigenetic Mechanisms - *CC BY-NC 4.0*
    - 2.6.2.5: Section 2 Final Evaluation - *CC BY-NC 4.0*
  - 2.6.3: Principles of Genetic Toxicology - *CC BY-NC 4.0*
    - 2.6.3.1: Introduction to Genetic-toxicology Assay - *CC BY-NC 4.0*
    - 2.6.3.2: Different Genetic Damages or Mutations - *CC BY-NC 4.0*
    - 2.6.3.3: Different Genetic-Toxicology Assays - *CC BY-NC 4.0*
    - 2.6.3.4: Different Cytotoxicity Assays - *CC BY-NC 4.0*
    - 2.6.3.5: Epigenetics Assay - *CC BY-NC 4.0*
    - 2.6.3.6: Section 3 Final Evaluation - *CC BY-NC 4.0*

- 2.6.4: Applied Systems Toxicology - *CC BY-NC 4.0*
  - 2.6.4.1: Systems Toxicology - *CC BY-NC 4.0*
  - 2.6.4.2: Dose Level and Applied Toxicology - *CC BY-NC 4.0*
  - 2.6.4.3: Tools and Technologies in Systems Toxicology - *CC BY-NC 4.0*
  - 2.6.4.4: Other Approaches for Predictive Toxicity Modeling - *CC BY-NC 4.0*
  - 2.6.4.5: Technologies Used In Systems Biology/Toxicology - *CC BY-NC 4.0*
  - 2.6.4.6: Takeaways Summary - *CC BY-NC 4.0*
  - 2.6.4.7: Section 4 Final Evaluation - *CC BY-NC 4.0*
- 2.6.5: Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.1: Introduction to Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.2: Global Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.3: Topic 3- Regional Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.4: National Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.5: State Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.6: Non-Governmental Regulatory Toxicology - *CC BY-NC 4.0*
  - 2.6.5.7: Section 5 Final Evaluation - *CC BY-NC 4.0*
- 2.6.6: Principles of Toxicology - *CC BY-NC 4.0*
  - 2.6.6.1: Introduction to Toxicology - *CC BY-NC 4.0*
    - 2.6.6.1.1: What is toxicology? - *CC BY-NC 4.0*
    - 2.6.6.1.2: Basic Terminology - *CC BY-NC 4.0*
  - 2.6.6.2: Dose and Dose Response - *CC BY-NC 4.0*
    - 2.6.6.2.1: Dose and It's Impact on Toxicity - *CC BY-NC 4.0*
    - 2.6.6.2.2: The Dose Response Relationship - *CC BY-NC 4.0*
    - 2.6.6.2.3: Dose Estimates of Toxic Effects - *CC BY-NC 4.0*
    - 2.6.6.2.4: Determining the Safety of a Drug - *CC BY-NC 4.0*
    - 2.6.6.2.5: NOAEL and LOAEL - *CC BY-NC 4.0*
  - 2.6.6.3: Toxic Effects - *CC BY-NC 4.0*
    - 2.6.6.3.1: Types of Toxic Effects - *CC BY-NC 4.0*
    - 2.6.6.3.2: Factors Affecting Toxicity - *CC BY-NC 4.0*
    - 2.6.6.3.3: Systemic Toxic Effects - *CC BY-NC 4.0*
    - 2.6.6.3.4: Organ Specific Toxic Effects - *CC BY-NC 4.0*
  - 2.6.6.4: Interactions - *CC BY-NC 4.0*
    - 2.6.6.4.1: Interactions - *CC BY-NC 4.0*
  - 2.6.6.5: Toxicity Testing Methods - *CC BY-NC 4.0*
    - 2.6.6.5.1: Testing and Assessing Toxicity - *CC BY-NC 4.0*
    - 2.6.6.5.2: Clinical Investigations and Other Types of Human Data - *CC BY-NC 4.0*
    - 2.6.6.5.3: Epidemiology Studies - *CC BY-NC 4.0*
  - 2.6.6.6: Risk Assessment - *CC BY-NC 4.0*
    - 2.6.6.6.1: Risk Assessment - *CC BY-NC 4.0*
    - 2.6.6.6.2: Hazard Identification - *CC BY-NC 4.0*
    - 2.6.6.6.3: Dose-Response Assessment - *CC BY-NC 4.0*
    - 2.6.6.6.4: Exposure Assessment - *CC BY-NC 4.0*
    - 2.6.6.6.5: Risk Characterization - *CC BY-NC 4.0*
  - 2.6.6.7: Exposure Standards and Guidelines - *CC BY-NC 4.0*
    - 2.6.6.7.1: Exposure Standards and Guidelines - *CC BY-NC 4.0*
    - 2.6.6.7.2: Regulation of Consumer Products and Drug Safety - *CC BY-NC 4.0*
    - 2.6.6.7.3: Environmental Exposure Standards/Guidelines - *CC BY-NC 4.0*
    - 2.6.6.7.4: Occupational (Workplace) Exposure Standards/Guidelines/Approaches - *CC BY-NC 4.0*
  - 2.6.6.8: Basic Physiology - *CC BY-NC 4.0*
    - 2.6.6.8.1: Introduction to Basic Physiology - *CC BY-NC 4.0*
    - 2.6.6.8.2: Homeostasis - *CC BY-NC 4.0*
    - 2.6.6.8.3: Organs and Organ Systems - *CC BY-NC 4.0*
    - 2.6.6.8.4: Tissues - *CC BY-NC 4.0*

- 2.6.6.8.5: Cells - *CC BY-NC 4.0*
- 2.6.6.8.6: Chemicals - *CC BY-NC 4.0*
- 2.6.6.9: Introduction to Toxicokinetics - *CC BY-NC 4.0*
  - 2.6.6.9.1: What is Toxicokinetics - *CC BY-NC 4.0*
- 2.6.6.10: Absorption - *CC BY-NC 4.0*
  - 2.6.6.10.1: Introduction to Absorption - *CC BY-NC 4.0*
  - 2.6.6.10.2: Gastrointestinal Tract - *CC BY-NC 4.0*
  - 2.6.6.10.3: Respiratory Tract - *CC BY-NC 4.0*
  - 2.6.6.10.4: Dermal Route - *CC BY-NC 4.0*
  - 2.6.6.10.5: Other Routes of Exposure - *CC BY-NC 4.0*
- 2.6.6.11: Distribution - *CC BY-NC 4.0*
  - 2.6.6.11.1: Introduction to Distribution - *CC BY-NC 4.0*
  - 2.6.6.11.2: Influence of Route of Exposure - *CC BY-NC 4.0*
  - 2.6.6.11.3: Disposition Models - *CC BY-NC 4.0*
  - 2.6.6.11.4: Structural Barriers to Distribution - *CC BY-NC 4.0*
  - 2.6.6.11.5: Storage Sites - *CC BY-NC 4.0*
- 2.6.6.12: Biotransformation - *CC BY-NC 4.0*
  - 2.6.6.12.1: Introduction to Biotransformation - *CC BY-NC 4.0*
  - 2.6.6.12.2: Chemical Reactions - *CC BY-NC 4.0*
  - 2.6.6.12.3: Biotransformation Sites - *CC BY-NC 4.0*
  - 2.6.6.12.4: Modifiers of Biotransformation - *CC BY-NC 4.0*
- 2.6.6.13: Excretion - *CC BY-NC 4.0*
  - 2.6.6.13.1: Introduction to Secretion - *CC BY-NC 4.0*
  - 2.6.6.13.2: Urinary Excretion - *CC BY-NC 4.0*
  - 2.6.6.13.3: Fecal Excretion - *CC BY-NC 4.0*
  - 2.6.6.13.4: Exhaled Air - *CC BY-NC 4.0*
  - 2.6.6.13.5: Other Routes - *CC BY-NC 4.0*
- 2.6.6.14: Cellular Toxicology - *CC BY-NC 4.0*
  - 2.6.6.14.1: Adaptation - *CC BY-NC 4.0*
  - 2.6.6.14.2: Cell Damage and Tissue Repair - *CC BY-NC 4.0*
  - 2.6.6.14.3: Cancer - *CC BY-NC 4.0*
  - 2.6.6.14.4: Neurotoxicity - *CC BY-NC 4.0*
- 2.6.6.15: Intuitive Toxicology and Risk Communication - *CC BY-NC 4.0*
  - 2.6.6.15.1: Intuitive Toxicology - *CC BY-NC 4.0*
  - 2.6.6.15.2: Risk Communication - *CC BY-NC 4.0*
- 2.6.6.16: Environmental Toxicology, Environmental Health, and One Health - *CC BY-NC 4.0*
  - 2.6.6.16.1: Environmental Toxicology - *CC BY-NC 4.0*
  - 2.6.6.16.2: Environmental Health - *CC BY-NC 4.0*
  - 2.6.6.16.3: One Health - *CC BY-NC 4.0*
- 2.6.6.17: Conclusion - *CC BY-NC 4.0*
- Back Matter - *CC BY-NC 4.0*
  - Detailed Licensing - *Undeclared*
- Back Matter - *Undeclared*
  - Index - *Undeclared*
  - Glossary - *Undeclared*
  - Detailed Licensing - *Undeclared*

3: Ecology - *CC BY-NC 4.0*

- 3.1: Ecology Overview - *CC BY-NC 4.0*
- 3.2: Populations - *CC BY-NC 4.0*
  - 3.2.1: Population Dispersion - *CC BY-NC 4.0*
  - 3.2.2: Population Size - *CC BY-NC 4.0*
  - 3.2.3: Population Growth and Regulation - *CC BY-SA 4.0*
  - 3.2.4: Life History - *CC BY-NC 4.0*
  - 3.2.5: Data Dive- Wildlife Corridors - *Undeclared*
  - 3.2.6: Review - *CC BY-NC 4.0*
- 3.3: Communities - *CC BY-NC 4.0*
  - 3.3.1: Biotic Interactions - *CC BY-NC 4.0*
    - 3.3.1.1: Trophic Interactions - *CC BY-NC 4.0*
      - 3.3.1.1.1: Predation - *CC BY-SA 4.0*
      - 3.3.1.1.2: Herbivory - *CC BY-SA 4.0*
      - 3.3.1.1.3: Parasitism - *CC BY-NC 4.0*
      - 3.3.1.1.4: Food Chains and Food Webs - *CC BY-NC 4.0*
    - 3.3.1.2: Competition - *CC BY-SA 4.0*
    - 3.3.1.3: Facilitation - *CC BY-NC 4.0*
    - 3.3.1.4: Symbiosis - *CC BY-NC 4.0*
  - 3.3.2: Community Structure and Dynamics - *CC BY-NC 4.0*

- 3.3.3: Data Dive- Bison Impacts In Prairies - *Undeclared*
- 3.3.4: Review - *CC BY-NC 4.0*
- 3.4: Ecosystems - *CC BY-NC 4.0*
  - 3.4.1: Ecosystem Types and Dynamics - *CC BY-NC 4.0*
  - 3.4.2: Matter - *CC BY-NC 4.0*
  - 3.4.3: Biogeochemical Cycles - *CC BY-SA 4.0*
  - 3.4.4: Soils - *CC BY-SA 4.0*
  - 3.4.5: Soil Degradation - *CC BY-NC-SA 4.0*
  - 3.4.6: Data Dive- Beaver Impacts on Wetlands - *Undeclared*
  - 3.4.7: Review - *CC BY-NC 4.0*
- 3.5: Biomes - *CC BY-SA 4.0*
  - 3.5.1: Climate and Biomes - *CC BY-NC-SA 4.0*
  - 3.5.2: Terrestrial Biomes - *CC BY-NC 4.0*
  - 3.5.3: Aquatic Biomes - *CC BY-SA 4.0*
  - 3.5.4: Data Dive- Biome Carbon Storage - *Undeclared*
  - 3.5.5: Review - *CC BY 4.0*

4: Conservation - *CC BY-NC 4.0*

- 4.1: The Value of Biodiversity - *CC BY-NC 4.0*
  - 4.1.1: Ecosystem Diversity - *CC BY-SA 4.0*
  - 4.1.2: Species Diversity - *CC BY-NC 4.0*
  - 4.1.3: Genetic Diversity - *CC BY-SA 4.0*
  - 4.1.4: Patterns of Biodiversity - *CC BY-SA 4.0*
  - 4.1.5: Data Dive- Biodiversity and Drugs - *Undeclared*
  - 4.1.6: Review - *CC BY-NC 4.0*
- 4.2: Threats to Biodiversity - *CC BY-NC 4.0*
  - 4.2.1: Extinction - *CC BY-NC 4.0*
  - 4.2.2: Measures of Biodiversity Loss - *CC BY-SA 4.0*
  - 4.2.3: Habitat Loss - *CC BY-SA 4.0*
  - 4.2.4: Overexploitation - *CC BY-SA 4.0*
  - 4.2.5: Pollution - *CC BY-SA 4.0*
  - 4.2.6: Invasive Species - *CC BY-NC 4.0*
  - 4.2.7: Climate Change - *CC BY-NC 4.0*
  - 4.2.8: Data Dive- Threats to Biodiversity - *Undeclared*
  - 4.2.9: Review - *CC BY-NC 4.0*
- 4.3: Protecting Biodiveristy - *CC BY-NC 4.0*
  - 4.3.1: Policies - *CC BY-NC 4.0*
  - 4.3.2: Non-profit Organizations - *CC BY-SA 4.0*
  - 4.3.3: Species-level Conservation - *CC BY-NC 4.0*
  - 4.3.4: Protected Areas - *CC BY-NC 4.0*
  - 4.3.5: Ecosystem Restoration - *CC BY-NC 4.0*
  - 4.3.6: Economic Influences on Conservation - *CC BY-SA 4.0*
  - 4.3.7: Individual Choices - *CC BY-NC 4.0*
  - 4.3.8: Data Dive- Island Fox Populations - *Undeclared*
  - 4.3.9: Review - *CC BY-NC 4.0*

5: Humans and the Environment - *CC BY-NC 4.0*

- 5.1: The Human Population - *CC BY-NC-SA 4.0*
  - 5.1.1: History of Human Population Growth - *CC BY-NC-SA 4.0*
  - 5.1.2: The Rate of Human Population Growth - *CC BY-SA 4.0*
  - 5.1.3: Looking Ahead - *CC BY-NC 4.0*
  - 5.1.4: Data Dive- World Population Densities - *Undeclared*
  - 5.1.5: Review - *CC BY-NC 4.0*
- 5.2: Water Resources - *CC BY-NC 4.0*
  - 5.2.1: Fresh Water Supply and the Water Cycle - *CC BY-NC-SA 4.0*
  - 5.2.2: Water Usage - *CC BY-NC-SA 4.0*
  - 5.2.3: Water Scarcity and Solutions - *CC BY-SA 4.0*
  - 5.2.4: Data Dive- Aqueducts of Rome - *Undeclared*
  - 5.2.5: Review - *CC BY-NC 4.0*
- 5.3: Agriculture - *CC BY-NC 4.0*
  - 5.3.1: Food Security - *CC BY-SA 4.0*
  - 5.3.2: Industrial Agriculture - *CC BY-SA 4.0*
  - 5.3.3: Selective Breeding and Genetic Engineering - *CC BY-SA 4.0*
  - 5.3.4: Sustainable Agriculture - *CC BY-SA 4.0*
  - 5.3.5: Data Dive- Landslides and Ag Terraces - *Undeclared*
  - 5.3.6: Review - *CC BY-NC 4.0*
- 5.4: Environmental Health - *CC BY 4.0*
  - 5.4.1: Types of Environmental Hazards - *CC BY-SA 4.0*
  - 5.4.2: Epidemiology - *CC BY-SA 4.0*
  - 5.4.3: Infectious Diseases - *CC BY-NC 4.0*
  - 5.4.4: Environmental Toxicology - *CC BY-SA 4.0*
  - 5.4.5: Environmental Hazard Reduction - *CC BY-SA 4.0*
  - 5.4.6: Data Dive- Cholera Cases Worldwide - *Undeclared*
  - 5.4.7: Review - *CC BY-NC 4.0*

6: Energy - *CC BY-NC 4.0*

- 6.1: Fossil Fuels - *CC BY 4.0*

- 6.1.1: Types of Fossil Fuels and Formation - *CC BY-NC-SA 4.0*
- 6.1.2: Mining, Processing, and Generating Electricity - *CC BY-NC-SA 4.0*
- 6.1.3: Fossil Fuel Consumption - *CC BY-NC 4.0*
- 6.1.4: Consequences of Fossil Fuels - *CC BY-NC-SA 4.0*
- 6.1.5: Data Dive- Global Fossil Fuel Consumption - *Undeclared*
- 6.1.6: Review - *CC BY-NC 4.0*

- 6.2: Nuclear Energy - *CC BY-SA 4.0*
  - 6.2.1: Radioactive Isotopes - *CC BY-NC 4.0*
  - 6.2.2: Generating Electricity with Nuclear Energy - *CC BY-NC 4.0*
  - 6.2.3: Nuclear Energy Consumption - *CC BY-NC 4.0*
  - 6.2.4: Consequences of Nuclear Energy - *CC BY-SA 4.0*
  - 6.2.5: Data Dive- Global Nuclear Power Generation - *Undeclared*
  - 6.2.6: Review - *CC BY-NC 4.0*
- 6.3: Renewable Energy - *CC BY-NC 4.0*
  - 6.3.1: Renewable Energy History and Consumption - *CC BY-SA 4.0*
  - 6.3.2: Wind Energy - *CC BY-NC 4.0*
  - 6.3.3: Solar Energy - *CC BY-NC 4.0*
  - 6.3.4: Geothermal Energy - *CC BY-NC 4.0*
  - 6.3.5: Hydropower - *CC BY-NC 4.0*
  - 6.3.6: Biofuels (Biomass Energy) - *CC BY-NC 4.0*
  - 6.3.7: Energy Conservation - *CC BY-NC 4.0*
  - 6.3.8: Data Dive- Global Renewable Energy Generation - *Undeclared*
  - 6.3.9: Review - *CC BY-NC 4.0*

- 7: Environmental Impacts - *CC BY 4.0*
  - 7.1: Solid Waste Management - *CC BY-SA 4.0*
    - 7.1.1: Waste Generation - *CC BY-NC 4.0*
    - 7.1.2: Waste Disposal - *CC BY-NC 4.0*
    - 7.1.3: Solid Waste and Marine Life - *CC BY-NC 4.0*
    - 7.1.4: Waste Reduction - *CC BY-NC 4.0*
    - 7.1.5: Data Dive- "Flushable" Wipes - *Undeclared*
    - 7.1.6: Review - *CC BY-NC 4.0*
  - 7.2: Pollution - *Undeclared*
    - 7.2.1: Water Pollution - *CC BY-NC 4.0*
      - 7.2.1.1: Water Pollutants and Their Sources - *CC BY-NC-SA 4.0*
      - 7.2.1.2: Water Treatment - *CC BY-SA 4.0*
      - 7.2.1.3: Mitigating Water Pollution - *CC BY-NC-SA 4.0*
    - 7.2.2: Air Pollution - *CC BY-SA 4.0*
      - 7.2.2.1: Ozone Depletion - *CC BY-SA 4.0*
      - 7.2.2.2: Acid Deposition - *CC BY 4.0*
    - 7.2.3: Data Dive- Plastic in World Oceans - *Undeclared*
    - 7.2.4: Review - *CC BY 4.0*
  - 7.3: Climate Change - *CC BY 4.0*
    - 7.3.1: The Greenhouse Effect and Climate Change - *CC BY-SA 4.0*
    - 7.3.2: Data Dive- Flooding in Venice Italy - *Undeclared*
    - 7.3.3: Review - *CC BY 4.0*
- 8: Sustainability - *CC BY-NC 4.0*
  - 8.1: Introduction to Sustainability - *Undeclared*
  - 8.2: History of Sustainability - *CC BY-NC 4.0*
  - 8.3: Environmental Ethics - *CC BY-NC 4.0*
    - 8.3.1: Environmental Justice and Indigenous Struggles - *CC BY-NC 4.0*
    - 8.3.2: Perspectives of the Environment - *CC BY-NC 4.0*
  - 8.4: Environmental Policy - *Undeclared*
  - 8.5: Sustainable Cities - *CC BY-NC 4.0*
  - 8.6: Sustainability and the Future - *CC BY-NC 4.0*
  - 8.7: Data Dive- Sustainable Investments - *Undeclared*
  - 8.8: Review - *CC BY 4.0*
- Back Matter - *Undeclared*
  - Index - *Undeclared*
  - Glossary - *Undeclared*
  - Detailed Licensing - *Undeclared*